ORSAY
Série A,
N° d'ordre :

204

THÈSES

PRESENTEES

UNIVERSITE DE PARIS-SUD
CENTRE D'ORSAY

POUR OBTENIR

LE TITRE DE DOCTEUR INGENIEUR

PAR

Jean-Patrice KLEIN

1ère THÈSE :

DÉTECTION DE CHANGEMENT DANS LES CARACTÉRISTIQUES D'UN PROCESSUS A REPRÉSENTATION GAUSSIENNE MARKOVIENNE

2ème THÈSE :

PROPOSITIONS DONNÉES

Soutenues le 20 décembre 1973 devant la Commission d'examen :

MM. PICINBONO		Président
BLAQUIERE	}	Examinateurs
FAURRE		

ORSAY
Série A,
N° d'ordre :

204

THÈSES

PRESENTEES

UNIVERSITE DE PARIS-SUD

CENTRE D'ORSAY

POUR OBTENIR

LE TITRE DE DOCTEUR INGENIEUR

PAR

Jean-Patrice KLEIN

1ère THÈSE :

DÉTECTION DE CHANGEMENT DANS LES CARACTÉRISTIQUES D'UN PROCESSUS A REPRÉSENTATION GAUSSIENNE MARKOVIENNE

2ème THÈSE :

PROPOSITIONS DONNÉES

Soutenues le 20 décembre 1973 devant la Commission d'examen :

MM. PICINBONO		Président
BLAQUIERE	}	Examinateurs
FAURRE	}	

UNIV.
LEIDEN
BIBL.

AVANT - PROPOS

Ce travail a été préparé à l'Institut de Recherche d'Informatique et d'Automatique (IRIA-LABORIA) .

Je voudrais exprimer toute ma reconnaissance à Monsieur le Professeur B. PICINBONO, dont les conseils et les remarques fort utiles m'ont permis de mener à bien cette thèse et qui m'a fait l'honneur d'accepter la présidence du jury d'examen .

Ma plus vive gratitude va à Monsieur P. FAURRE, Directeur Scientifique à l'IRIA pour l'aide et les encouragements qu'il m'a prodigués et auxquels ce travail doit beaucoup.

Je remercie beaucoup Monsieur le Professeur BLAQUIERE d'avoir bien voulu faire partie du Jury.

Merci enfin à Madame BARNY qui s'est chargée de dactylographier le manuscrit, ainsi qu'à tous mes camarades de l'IRIA, en particulier MM. ATTASI, BERNHARD, CLERGET, GERMAIN, LEVIEUX, pour leur aide .

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41542716_0016

SOMMAIRE

	Page
Notations	
Chapitre I - Introduction	1
Chapitre II - Rappels : Théorie de la détection, application aux modèles gaussiens markoviens	3
Chapitre III - Détection de changement brusque dans les caractéristiques d'un processus à représentation gaussienne-markovienne	19
Chapitre IV - Etude du comportement du test de détection entre deux structures gaussiennes markoviennes dans le cas général	34
Chapitre V - Détection de changement des caractéristiques d'un processus à représentation gaussienne markovienne	54
Chapitre VI - Etude expérimentale des algorithmes proposés.	65
Chapitre VII - Conclusion générale	79
Bibliographie	81
Appendices	83
Annexes	

Table des Notations

F, H, R, Q	:	matrices
F', K'	:	transposée de la matrice ou du vecteur
$X(t)$	:	processus stochastique vectoriel
$y(t)$	:	processus stochastique scalaire
$\mathcal{Y}(t)$	:	$\{y(\tau),\ \tau \in [0,t]\}$
$\mathcal{Y}_n$	:	$\{y_1, y_2, \ldots y_n\}$
Δt	:	période d'échantillonnage
a	:	$\{a^0, a^1, \ldots a^l\}$, ensemble des paramètres sur lesquels portent les hypothèses
a_o	:	$\{a_o^0, a_o^1, \ldots a_o^l\}$
H_o, H_1	:	hypothèses simples portant sur la valeur de a ($a = a_o$ ou $a = a_1$)
$\mathcal{H}_o, \mathcal{H}_1$	:	" composées ou multiples portant sur la valeur de a
α, β	:	probabilités d'erreur de première et seconde espèce
$\mathcal{T}_i$	:	equation du filtre de Kalman utilisant le modèle markovien de paramètres d'indice i ($i = o, 1$)
[9]	:	référence n° 9 dans la liste bibliographique
[V]	:	appendice N° V

Chapitre I

Introduction

La nécessité d'analyser et de prévoir le comportement des systèmes, a conduit à modéliser les processus aléatoires mis en jeu par ces systèmes. Dans la mesure où les processus considérés sont gaussiens, (une bonne partie des processus rencontrés dans la pratique le sont du fait qu'ils sont générés par "addition" de perturbations aléatoires indépendantes) et à spectre rationnel (ou approximable par une fonction rationnelle), on peut en construire un modèle linéaire. On sait alors, que la résolution des problèmes de filtrage et de commande revient principalement à l'intégration d'une équation de type Riccati. Une telle représentation des processus aléatoires par équations différentielles linéaires et stochastiques est appelée représentation gaussienne markovienne. Une étude des propriétés de ces représentations est présentée dans FAURRE [5].

Il est commode de considérer des modèles stationnaires, (caractéristiques invariantes relativement à une translation dans le temps) bien qu'en pratique cette hypothèse ne soit pas constamment vérifiée :
Les caractéristiques du modèle peuvent être modifiées par suite de causes telles que :

- panne ou erreur survenue dans le système physique observé par le processus
- variations affectant les constantes du système physique
- présence ou absence d'un signal dans un bruit (écho radar, etc...).

Le problème que nous nous proposons de traiter ici est le suivant : Disposant d'un modèle mathématique du processus $y(t)$, on cherche à détecter un changement dans les caractéristiques de ce modèle.
Ce problème relève de la théorie des tests d'hypothèses.

Les processus y(t) considérés ici, se limitent à la classe des processus à représentation gaussienne-markovienne stationnaires. En outre, ces processus sont scalaires et discrets.

La plupart des publications sur les problèmes de détection supposent connu le changement susceptible d'intervenir sur les caractéristiques du modèle de Y(t). (voir NEWBOLD et HO [10], LUMELSKII [9]).
Deux modèles seulement sont possibles, chaque hypothèse portant sur un seul modèle (hypothèse "simple"). Mais dans la pratique, on ignore généralement la valeur exacte des paramètres après changement. On sait seulement que le modèle final appartient à un certain ensemble de modèles.
De "simples" les deux hypothèses deviennent "multiples", en ce sens qu'elles portent sur plusieurs modèles possibles.

Dans cette étude, nous aborderons le problème de détection de changement de caractéristiques dans le modèle de y(t), la valeur du changement n'étant pas connue.

Après rappel de certaines notions, sur la théorie de la détection (en particulier sur la détection en deux structures gaussiennes markoviennes), nous traiterons de la détection de changement brusque au changement par saut, le changement étant alors connu (chapitres II et III).

Le quatrième chapitre analysera la sensibilité de la détection entre deux structures gaussiennes markoviennes aux erreurs de modèles.

Le cinquième chapitre étudiera la détection de changement brusque et la détection de variation des caractéristiques de y(t), le changement et la variation n'étant pas connue. Deux algorithmes sont proposés.

Le dernier chapitre est consacré à l'étude expérimentale de ces deux algorithmes.

Chapitre II

Rappels ; Théorie de la détection, application aux modèles gaussiens markoviens

Ce chapitre a pour but de rassembler des notions et résultats qui serviront dans cette étude.
Dans un premier paragraphe, nous présentons le problème de la détection et les façons de le résoudre : détection globale et détection séquentielle.
Puis, nous exposons le test séquentiel de Wald que nous utiliserons pour les problèmes traités dans les chapitres suivants. On pourra consulter les publications de Wald [16], [15], Bussgang et Middleton [2], Hancock et Wintz [6] pour préciser les aspects statistiques de ce test.
Enfin, nous présentons l'application de la détection séquentielle aux modèles gaussiens markoviens.
Les processus étudiés étant discrets, nous ne considérons que des observations discrètes.

II. 1. Problème de la détection.

Soient deux hypothèses H_o et H_1 portant sur un système physique, observable par un processus stochastique $y(t)$.
Le problème est le suivant :

On désire pouvoir décider de façon optimale, en un sens que nous allons préciser, si l'on se trouve dans l'hypothèse H_o ou dans l'hypothèse H_1 au moyen de l'observation $\mathcal{Y}(t_n)$: $\mathcal{Y}(t_n) = \{y(t_o), y(t_1), \ldots, y(t_{n-1}), y(t_n)\}$.

II. 1. 1. Détection globale.

N nombre d'observations est fixé.
Soit E_N, l'espace formé de toutes les réalisations possibles de $\mathcal{Y}(t_N)$.

On résoud le problème en scindant l'espace E_N en deux sous-espaces E_0 et E_1 tels que :

si $\mathcal{Y}(t_N) \in E_o$, on accepte H_o.

si $\mathcal{Y}(t_N) \in E_1$, on accepte H_1.

L'optimalité intervient dans le choix de E_o, E_1.

Introduisons les coûts suivants :

C_{ij} = coût associé à l'acceptation de l'hypothèse H_i, quand l'hypothèse H_j est vraie ($i = 0,1$; $j = 0,1$)

On suppose connues les probabilités à priori, $p(H_1)$ et $p(H_o)$.

On définit alors le critère suivant, appelé risque Bayesien :

$$(2\text{-}1) \qquad R = C_{oo}\, p(H_o) \int_{E_o} p[\mathcal{Y}_N/H_o]\, d\mathcal{Y}_N + C_{o1}\, p(H_1) \cdot \int_{E_o} p[\mathcal{Y}_N/H_1]\, d\mathcal{Y}_N$$

$$+ C_{1o}\, p(H_o) \int_{E_1} p[\mathcal{Y}_N/H_o]\, d\mathcal{Y}_N + C_{11}\, p(H_1) \int_{E_1} p[\mathcal{Y}_N/H_1]\, d\mathcal{Y}_N$$

E_0 et E_1 vont être choisis de façon à minimiser ce critère.

1°) Choix de E_o, E_1.

Pour la commodité de la démonstration, posons :

$C_\alpha = C_{10} - C_{oo} > o$.

$C_\beta = C_{01} - C_{11} > o$.

L'hypothèse $C_\alpha > o$ et $C_\beta > o$ est satisfaisante dans n'importe quelle situation, le coût associé à une erreur étant plus élevé que le coût associé à une décision correcte.

Dans beaucoup d'applications, on pose : $C_{oo} = C_{11} = 0$

Compte tenu de l'hypothèse précédente :

$$(2\text{-}2)\quad R = C_{oo}p(H_o) + C_{11}p(H_1) + C_\alpha p(H_o)\int_{E_1} p[y_N/H_o]dy_N + C_\beta p(H_1)\cdot\int_{E_o} p[y_N/H_1]dy_n.$$

Or $$\int_E p[y_N/H_o]dy_N = 1 = \int_{E_1} p[y_N/H_o]dy_N + \int_{E_o} p[y_N/H_o]\cdot dy_N$$

Le risque Bayesien s'écrit alors :

$$(2\text{-}3)\quad R = C_{oo}p(H_o) + C_{11}p(H_1) + C_\alpha p(H_o) + \int_{E_o}\Big[C_\beta p(H_1)p[y_N/H_1] - C_\alpha p(H_o)p[y_N/H_o]\Big]dy_N$$

Les trois premiers termes étant constants, R est minimum quand l'intégrale est minimum

Comme : $C_\beta p(H_1)\, p[y_N/H_1] > o$

$C_\alpha p(H_o)\, p[y_N/H_o] > o$,

alors toutes les valeurs y_N telles que :

$$C_\beta\, p(H_1)\, p[y_N/H_1] > C_\alpha\, p(H_o)p[y_N/H_o]$$

contribuent à faire croître R, l'intégrande de (2-3) étant positif. Elles doivent donc appartenir à E_1.

De même, toutes les valeurs y_N telles que :

$$C_\alpha\, p(H_o)\, p[y_N/H_o] > C_\beta\, p(H_1)\, p[y_N/H_1]$$

contribuent à faire décroître R, l'intégrande de (2-3) étant négatif. Ces valeurs appartiennent donc à E_o.

E_o et E_1 sont donc définis par :

$$y_N \in E_o \Leftrightarrow \frac{p[y_N/H_1]}{p[y_N/H_o]} < \frac{C_\alpha\, p(H_o)}{C_\beta\, p(H_1)}$$

$$\mathcal{Y}_N \in E_1 \Leftrightarrow \frac{p[\mathcal{Y}_N/H_1]}{p[\mathcal{Y}_N/H_o]} > \frac{C_\alpha\, p(H_o)}{C_\beta\, p(H_1)} .$$

2°) Test de Bayes.

Posons : $$\Lambda(\mathcal{Y}_N) = \frac{p[\mathcal{Y}_N/H_1]}{p[\mathcal{Y}_N/H_o]}$$

$$K = \frac{C_\alpha\, p(H_o)}{C_\beta\, p(H_1)}$$

$\Lambda(\mathcal{Y}_N)$ est appelé rapport de vraisemblance.
K est appelé le seuil.

Le test de Bayes est alors le suivant :

Etant donné une observation $\mathcal{Y}_N = \mathcal{Y}^*_N$, on calcule $\Lambda(\mathcal{Y}^*_N)$ et on le compare au seuil K ;

si $\Lambda(\mathcal{Y}^*_N) > K$ on choisit l'hypothèse H_1.

si $\Lambda(\mathcal{Y}^*_N) < K$ on choisit l'hypothèse H_o.

La séparation de E en deux domaines E_o et E_1, provient de considérations statistiques. Il s'ensuit une certaine probabilité d'erreur.

Soient :

α, probabilité de rejeter H_o quand H_o est vraie.

β, probabilité d'accepter H_o quand H_1 est vraie,

nous avons les relations :

(2-4)
$$\alpha = \int_{E_1} p[\mathcal{Y}_N/H_o]\, d\mathcal{Y}_N$$

$$\beta = \int_{E_o} p[\mathcal{Y}_N/H_1]\, d\mathcal{Y}_N$$

α et β sont respectivement appelées probabilité d'erreur de première espèce et probabilité d'erreur de seconde espèce.

La solution de Bayes au problème de test d'hypothèses statistiques est une stratégie optimale ; en ce sens, qu'aucune autre stratégie ne peut atteindre un risque moyen plus petit.

II.1.2. Détection séquentielle

Sous l'hypothèse du nombre d'observations fixé à l'avance, le test de Bayes est optimum.

Mais si cette contrainte du nombre fixé d'observations est levée, il est possible de construire des tests d'hypothèses qui pour les mêmes probabilités d'erreurs, sont supérieurs au test de Bayes, en ce sens que le nombre moyen d'observations est plus faible.

Cependant, ces tests ont le désavantage suivant : le nombre d'observations nécessaire pour détecter est une variable aléatoire. Dans le paragraphe suivant, nous présentons une méthode de détection séquentielle. Ce test fut construit par Wald [15].
Nous appelons puissance d'un test, le couple (α,β).
R défini en (2-2) peut s'écrire :

$$R = C_{oo}\, p(H_o) + C_{11}\, p(H_1) + C_{\alpha}\, p(H_o).\, \alpha + C_{\beta}\, p(H_1).\, \beta$$

Pour des tests de même puissance, R est le même.

Pour comparer des tests de détection de même puissance, mais avec un nombre d'observations différentes, nous devons faire intervenir le coût des observations.

Si toutes les observations sont de coût égal, on peut définir un risque moyen :

$$R_S = R + C.\, \bar{n} \qquad (2\text{-}5)$$

où R est le risque moyen défini par (2-2), C le coût d'une observation, et $\bar{n}$ le nombre moyen d'observations nécessaire pour détecter.

Le test de Wald minimise le risque moyen R_S.

Wald et Wolfowitz ont montré [16] que, étant données les probabilités d'erreur (α,β), les coûts (C_α, C_β, C), les probabilités a priori $p(H_o)$ et $p(H_1)$, aucun autre test ne nécessite un nombre moyen d'observations plus faible que le test de Wald.

II.2. Test de Wald

Nous considérons toujours deux hypothèses H_o et H_1 portant sur la structure du processus stochastique $y(t)$. (Les hypothèses peuvent porter sur la valeur d'un ou plusieurs paramètres du modèle de $y(t)$).

Soit : $a = \{a^1, a^2, \ldots a^l\}$ l'ensemble des paramètres sur lesquels portent les hypothèses.

Les hypothèses sont les suivantes :

H_o : $a = a_o$ $\qquad a_o = \{a_o^1, a_o^2, \ldots a_o^l\}$.

H_1 : $a = a_1$ $\qquad a_1 = \{a_1^1, a_1^2, \ldots a_1^l\}$

Un test séquentiel se fait par étapes successives.

A chaque étape, l'espace des observations E_n est divisé en trois zones :

E_{1n}, E_{on} et E'_n .

Si $\underline{y}_n \in E_{1n}$, le test s'arrête avec acceptation de l'hypothèse H_1.

Si $\underline{y}_n \in E_{on}$, le test s'arrête avec acceptation de l'hypothèse H_o.

Si $\underline{y}_n \in E'_n$, le test continue en prenant une nouvelle observation $y(t_{n+1})$.

La zone E'_n est appelée zone d'indifférence.

Regardons comment se fait la division de l'espace des observations dans le cas du test de Wald.

II. 2. 1. Procédure du test de Wald.

Considérons le rapport de vraisemblance, $\Lambda(\mathcal{Y}_n)$:

$$\Lambda(\mathcal{Y}_n) = \frac{p[\mathcal{Y}_n/H_1]}{p[\mathcal{Y}_n/H_o]}$$

La division de l'espace des observations se fait en sélectionnant deux nombres, A et B et en établissant les règles de procédure suivantes :

à la n^e étape :

si $\Lambda(\mathcal{Y}_n) \geqslant A$, $\mathcal{Y}_n \in E_{1n}$ et l'hypothèse H_1 est acceptée.

si $\Lambda(\mathcal{Y}_n) \leqslant B$, $\mathcal{Y}_n \in E_{on}$ et l'hypothèse H_o est acceptée.

si $B < \Lambda(\mathcal{Y}_n) < A$, $\mathcal{Y}_n \in E'_n$, on continue le test.

On continue cette procédure jusqu'à ce que H_o ou H_1 soit acceptée.

Nous pouvons également comparer n'importe quelle fonction monotone de ces variables $(\Lambda(\mathcal{Y}_n), A, B)$.

La plus utilisée, est la fonction logarithme, $\Lambda(\mathcal{Y}_n)$ étant dans beaucoup de cas un rapport de probabilités s'exprimant sous forme d'exponentielles.

Soit : $\theta_n = \text{Log } \Lambda(\mathcal{Y}_n)$

La procédure est alors la suivante :

à la n^e étape, on calcule θ_n ;

si $\theta_n \geqslant$ Log A, H_1 est acceptée

si $\theta_n \leqslant$ Log B, H_o est acceptée

si Log B $< \theta_n <$ Log A, on prend une nouvelle observation.

Une question importante est de savoir si le test se terminera.

Stein [13] a montré que tous les moments de n sont finis si les observations sont indépendantes.

En outre, Wald [15] indique que le test se termine presque surement par une large classe de densité de probabilité, même quand les observations sont correlées.

Il reste maintenant à savoir, comment fixer les bornes A et B, et quelles sont les probabilités d'erreur correspondantes au test.

II. 2. 2. Relations entre les bornes A,B et les probabilités d'erreur α,β.

Supposons qu'à la m^e étape, le test se termine par l'acceptation de H_1

Il vient : $\Lambda(\mathcal{Y}_n) \geqslant A$

ou : (2-6) $p[\mathcal{Y}_m/H_1] \geqslant A\, p[\mathcal{Y}_m/H_o]$

Soit E', l'ensemble des réalisations de $\mathcal{Y}_m$ conduisant à l'acceptation de H_1.

(2-6) est équivalent à :

$$(2\text{-}7) \qquad \int_{E'} p[\mathcal{Y}_m/H_1]\, d\mathcal{Y}_m \geqslant A \int_{E'} p[\mathcal{Y}_m/H_o]\, d\mathcal{Y}_m .$$

Or, par définition, l'intégrale de gauche est : $1 - \beta$, l'intégrale de droite est α

Alors : $1 - \beta \geqslant A\alpha$

De façon similaire, on obtient : $\beta \leqslant B(1-\alpha)$

A et B sont donc donnés par :

$$(2\text{-}8) \qquad A \leqslant \frac{1-\beta}{\alpha}$$

$$B \geqslant \frac{\beta}{1-\alpha}$$

Par hypothèse, on considère habituellement que les dépassements de $\Lambda(\mathcal{Y}_m)$ au dessus des bornes A, B sont négligeables. (par exemple, période d'échantillonnage peu élevée).

On peut alors écrire :

$$(2\text{-}9) \qquad \begin{cases} A = \dfrac{1-\beta}{\alpha} \\ \\ B = \dfrac{\beta}{1-\alpha} \end{cases}$$

Remarque :

Dans le cas d'observations continues, la fonction $\Lambda(\mathcal{Y}(t))$ est continue en t (voir Bussgang et Middleton [2]).

Les relations (2-9) deviennent alors exactes.

II. 2. 3. Performances d'un test séquentiel.

Le test d'hypothèses H_o, H_1 suppose que la valeur du paramètre a est a_o ou a_1.

Mais dans bien des cas, la vraie valeur de a n'est pas connue.

Il est alors utile de savoir comment se comporte le test d'hypothèses H_o, H_1, quelles sont les performances quand la vraie valeur de a n'est ni a_o , ni a_1.

Deux fonctions permettent d'évaluer ces performances :

L(a) : probabilité pour que le test se termine avec l'acceptation de H_o, la vraie valeur du paramètre étant a.

E[n(a)] : nombre moyen d'observations nécessaire pour prendre une décision, la vraie valeur du paramètre étant a.

1°) fonction L(a).

On montre (voir appendice [I])que sous certaines hypothèses, L(a) peut se mettre sous la forme :

$$L(a) = \frac{A^h - 1}{A^h - B^h} \qquad (2\text{-}10)$$

où h est tel que : $- h \neq o$

$- E[A^h(\mathcal{Y}_m)] = 1$

2°) fonction E[n(a)]

Nous considérons ici le cas d'observations discrètes.

Soit $E_a[\theta_n/n = N]$ la valeur moyenne de θ_n sachant que :

- le test se termine à l'étape N
- la vraie valeur du paramètre est a.

N, point d'arrêt du test est une variable aléatoire dépendant de a.

Posons : $N = n(a)$.

Compte tenu de la définition de N, θ_N ne peut prendre que deux valeurs : Log A ou Log B avec des probabilités respectives de 1 - L(a) et L(a).

On a donc :

$$E_a[\theta_n/n = N] = L(a) \text{ Log } B + (1 - L(a)) \text{ Log } A \qquad (2\text{-}11)$$

Deux cas se présentent : les observations sont indépendantes ou les observations sont correlées.

- observations indépendantes

On peut écrire : $P[\mathcal{Y}_n/H] = \prod_{i=1}^{n} p[y_i/H]$

Posons : $Z(a) = \text{Log} \dfrac{p[y/a_1]}{p[y/a_o]}$

on obtient pour n fixe : $E_a[\theta_n] = n\, E[Z(a)]$

d'où : $E_a[\theta_n/n = N] = E[n(a)]\, E[Z(A)]$

et : (2-12) $E[n(a)] = \dfrac{L(a)\ \text{Log}\ B + (1-L(a))\text{Log}\ A}{E[Z(a)]}$

- observations correlées

Dans le cas d'observations correlées, il est encore possible d'obtenir une relation simple pour $E[n(a)]$, à condition que $E_a[\theta_n]$ puisse s'exprimer comme fonction linéaire de n (voir Bussgang et Middleton [2]).

II. 2. 4. Troncature du test séquentiel.

Le nombre d'observations nécessaire pour détecter n, étant une variable aléatoire, la probabilité pour que n soit grand est non nulle.

Si la valeur du paramètre a se trouve entre a_o et a_1, et si les probabilités d'erreur α et β sont petites, le nombre moyen d'observations nécessaire pour détecter peut également devenir très grand.

Ces deux raisons peuvent rendre un test séquentiel non satisfaisant.

Dans certaines situations, il est alors préférable d'interrompre le test et de décider en faveur de telle ou telle hypothèse.

Cette procédure de test est appelée test séquentiel tronqué.

Le test séquentiel normal est mis en oeuvre jusqu'à ce qu'une décision soit prise ou que l'étape N soit atteinte.

A l'étape N, si aucune décision n'a été prise,

on accepte l'hypothèse H_o si $\theta_N < o$

" " H_1 si $\theta_N > o$

Le test séquentiel tronqué se termine donc au plus en N étapes.

Un test séquentiel tronqué est un compromis entre un test séquentiel normal et un test de Bayes classique, où le nombre d'observations est fixé.

II. 2. 5. Tests séquentiels d'hypothèses multiples.

Nous venons de décrire la méthode séquentielle de Wald pour tester deux hypothèses simples H_o et H_1. Nous avons été capable de déterminer les performances du test, la vraie valeur du paramètre pouvant être différente de a_o et a_1.

Nous présentons ici une généralisation du test de Wald, nous permettant de considérer des hypothèses multiples.

1°) Préliminaires

Soient les deux hypothèses composées,

$\mathcal{H}_o$: a peut prendre une valeur appartenant à un ensemble Ω_o.

$\mathcal{H}_1$: a " " " " Ω_1.

Nous supposons :

- $\Omega_o \cap \Omega_1 = \emptyset$

- il est possible de définir deux fonctions non négatives $p_o(a)$ et $p_1(a)$ définies respectivement sur Ω_o et Ω_1, telles que : $\int_{\Omega_o} p_o(a) = 1$

$$\int_{\Omega_1} p_1(a) = 1$$

$p_o(a)$ et $p_1(a)$ peuvent être interprêtées comme des densités de probabilités sur Ω_o et Ω_1.

2°) Procédure du test

La procédure du test séquentiel des hypothèses $\mathcal{H}_0$, $\mathcal{H}_1$ est alors la suivante :

A la n° étape,

- on calcule le rapport de vraisemblance généralisé :

$$\Lambda(\mathcal{Y}_n) = \frac{\int_{\Omega_1} p_1(a)\, p[\mathcal{Y}_n/a]\, da}{\int_{\Omega_0} p_0(a)\, p[\mathcal{Y}_n/a]\, da} \qquad (2\text{-}13)$$

- on le compare à deux bornes A et B :

si $\Lambda(\mathcal{Y}_n) > A$ on accepte $\mathcal{H}_1$

si $\Lambda(\mathcal{Y}_n) < B$ " $\mathcal{H}_0$

si $B < \Lambda(\mathcal{Y}_n) < A$ on prend une nouvelle observation.

3°) Relation entre les bornes et les probabilités d'erreurs.

Par des considérations identiques au test d'hypothèses simples, on choisit :

$$A = \frac{1-\bar{\beta}}{\bar{\alpha}}$$

$$B = \frac{\bar{\beta}}{1-\bar{\alpha}}$$

où $\bar{\alpha}$ et $\bar{\beta}$ sont les probabilités d'erreur moyennes définies par :

$$\begin{cases} \bar{\alpha} = \int_{\Omega_0} \alpha(a)\, p_0(a)\, da \\ \bar{\beta} = \int_{\Omega_1} \beta(a)\, p_1(a)\, da \end{cases} \qquad (2\text{-}14)$$

La procédure du test séquentiel d'hypothèses composées (ou d'hypothèses multiples), est identique à celle du test d'hypothèses simples, à la seule différence que l'on se sert du rapport de vraisemblance généralisé et de probabilités d'erreur moyennes.

II. 3. Détection séquentielle entre deux modèles à représentation gaussienne markovienne.

Le processus stochastique y(t) est à représentation gaussienne markovienne. Nous voulons détecter le modèle de y(t) parmi deux modèles donnés. Nous utiliserons le test séquentiel de Wald du fait de son caractère d'optimalité précisé dans les paragraphes précédents.
La solution exposée est celle de Schweppe [12].

II. 3. 1. Processus à représentation gaussienne markovienne.

Soit y(t) un processus gaussien centré stationnaire.

Un tel processus sera dit à représentation gaussienne markovienne s'il peut être mis sous la forme :

$$(2\text{-}15) \qquad \begin{aligned} \dot{X}(t) &= F\,X(t) + v(t) \\ y(t) &= H\,X(t) + w(t) \end{aligned}$$

où $\begin{bmatrix} v(t) \\ w(t) \end{bmatrix}$ est un bruit blanc gaussien centré de covariance :

$$E\left[\begin{bmatrix} v(t) \\ w(t) \end{bmatrix} [v'(s), w'(s)]\right] = \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \delta(t-s)$$

De plus, les conditions suivantes doivent être vérifiées :

F = matrice asymptotiquement stable.
(F,L) = paire complètement commandable avec Q = LL'
(H,F) = paire complètement observable.

Ces processus sont étudiés dans FAURRE [5].

II. 3. 2. Détermination du rapport de vraisemblance.

Les deux hypothèses à tester sont :

H_o ; y(t) a pour modèle (o)
$$\begin{cases} \dot{X}_o(t) = F_o\,X_o(t) + v_o(t) \\ y(t_n) = H\,X_o(t_n) + w_o(t_n) \end{cases}$$

H_1 : $y(t)$ a pour modèle (1)

$$\begin{cases} \dot{X}_1(t) = F_1 X_1(t) + v_1(t) \\ y(t_n) = H X_1(t_n) + w_1(t_n) \end{cases}$$

Les observations sont discrètes.

Le rapport de vraisemblance est : $\Lambda(\mathcal{Y}_n) = \dfrac{p[\mathcal{Y}_n/H_1]}{p[\mathcal{Y}_n/H_o]}$

sachant que $p[\mathcal{Y}_n/H_i] = p[y(t_n)/\mathcal{Y}_{n-1}, H_i] \cdot p[\mathcal{Y}_{n-1}/H_i]$ $(i = o,1)$

$\Lambda(\mathcal{Y}_n)$ peut aussi s'écrire :

$$\Lambda(\mathcal{Y}_n) = \Lambda(\mathcal{Y}_{n-1}) \cdot \frac{p[y(t_n)/H_1, \mathcal{Y}_{n-1}]}{p[y(t_n)/H_o, \mathcal{Y}_{n-1}]}$$

ou encore, en posant $\theta_n = \text{Log}\,[\Lambda(\mathcal{Y}_n)]$:

$$\theta_n = \theta_{n-1} + \text{Log}\,\frac{p[y(t_n)/\mathcal{Y}_{n-1}, H_1]}{p[y(t_n)/\mathcal{Y}_{n-1}\ H_o]} \qquad (2\text{-}16)$$

Il suffit alors de calculer $p[y(t_n)/\mathcal{Y}_{n-1}, H_i]$ $(i = 0,1)$ pour calculer θ_n.

1°) Calcul de $p[y(t_n)/\mathcal{Y}_{n-1}, H_i]$

L'estimation de l'état $X(t)$ à l'aide du filtre de Kalman, va nous permettre de calculer l'expression de $p[y(t_n)/\mathcal{Y}_{n-1}, H_i]$.

En effet sous l'hypothèse H_i (modèle i, i = 0,1) le filtre de Kalman $\mathcal{F}_i$ donne l'expression de :

$$E[y(t_n)/\mathcal{Y}_{n-1}, H_i] = \hat{y}_i(t_n)$$

$$\text{cov}\,[y(t_n)/\mathcal{Y}_{n-1}, H_i] = \sigma_i^2(t_n)$$

Les équations donnant $\hat{y}_i(t_n)$ et $\sigma_i^2(t_n)$ sont exposées en appendice [II].

Quand le filtre de Kalman a atteint son régime stationnaire, alors :

$$\sigma_i^2(t_n) = \sigma_{i,\infty}^2 = \sigma_i^2$$

L'expression de la densité de probabilité est alors :

$$p[y(t_n)/\mathcal{Y}_{n-1}, H_i] = \frac{1}{\sigma_i \sqrt{2\pi}} \; e^{-\frac{1}{2} \frac{(y(t_n)-\hat{y}(t_n))^2}{\sigma_i^2}} \qquad (2\text{-}17)$$

2°) Calcul de θ_n.

On utilise deux filtres, $\mathcal{J}_o$ et $\mathcal{J}_1$ basés respectivement sur les modèles (o) et (1)

Posons : $\varepsilon_{in} = y(t_n) - \hat{y}_i(t_n)$

Compte tenu de (2-16) et (2-17), l'expression de θ_n devient :

$$\theta_n = \theta_{n-1} + \text{Log} \frac{\sigma_o}{\sigma_1} - \frac{1}{2}\left[\frac{\varepsilon_{1n}^2}{\sigma_1^2} - \frac{\varepsilon_{on}^2}{\sigma_o^2}\right] \qquad (2\text{-}18)$$

II. 3. 3. Algorithme de détection séquentielle.

On se fixe les probabilités d'erreur α et β, d'où on déduit les bornes A et B à l'aide de la relation (2-9).

L'algorithme est le suivant :

1 - prendre une nouvelle mesure $y(t_n)$

2 - estimation de l'état avec les hypothèses H_o et H_1.

3 - en déduire ε_{on} et ε_{1n}.

4 - calculer θ_n donné par (2-18)

5 - Comparer θ_n à Log A et Log B :

si Log B $< \theta_n <$ Log A, retourner en 1.

si $\theta_n \geqslant$ Log A, accepter H_1

si $\theta_n \leqslant$ Log B, accepter H_o.

II. 4. Conclusion

La détection séquentielle présente plusieurs avantages par rapport à la détection globale, en particulier :

- Pour les mêmes probabilités d'erreur, le nombre d'observations nécessaire pour prendre une décision est en moyenne plus faible.

- Les observations sont traitées au fur et à mesure de leur arrivée. Cette méthode de détection est donc bien adaptée au temps réel. Elle ne nécessite qu'un faible volume de mémoire.

Le principal inconvénient est que l'on ne connait pas le nombre exact d'observations nécessaire pour détecter. Ce nombre est une variable aléatoire, il peut être très petit comme très grand, d'où la nécessité de tronquer parfois le test.

Nous avons présenté une méthode de détection séquentielle : le test de Wald, et souligné son caractère d'optimalité par rapport aux autres tests séquentiels.
Dans la suite, nous utiliserons cette méthode de test.

Cette étude se fait dans le cadre des processus à représentation gaussienne markovienne. Dans le dernier paragraphe, nous avons souligné les liaisons entre les problèmes de filtrage et de détection. En particulier, les équations du filtre de Kalman, permettent de trouver une solution récursive simple pour la détermination du rapport de vraisemblance.

Chapitre III

Détection de changement brusque dans les caractéristiques d'un processus à représentation gaussienne-markovienne (le saut étant connu)

Soit $y(t)$, un processus à représentation gaussienne-markovienne. A l'instant $t = o$, le modèle de $y(t)$ est (o) :

$$(o) \quad \begin{cases} \dot{X}_o(t) = F_o\, X_o(t) + v_o(t) \\ y(t_n) = H\, X_o(t_n) + w(t_n) \end{cases}$$

Le modèle de $y(t)$ est susceptible de passer brusquement de (o) à (1) à un instant τ :

$$(1) \quad \begin{cases} \dot{X}_1(t) = F_1\, X_1(t) + v_1(t) \\ y(t_n) = H\, X_1(t_n) + w(t_n) \end{cases}$$

Le problème posé est celui de détecter le moment τ, du changement brusque des caractéristiques de $y(t)$, aussi rapidement que possible, tout en imposant une probabilité d'erreur assez faible.

Bien qu'apparenté, ce problème est distinct de celui présenté dans le chapitre précédent où l'on cherchait à détecter si les caractéristiques du modèle appartenaient à un ensemble hypothétique de valeurs par rapport à un autre ensemble. Ici, on cherche à détecter le moment du changement aussi rapidement que possible.

La détection de changement brusque de paramètres a été l'objet de plusieurs publications parmi lesquelles nous mentionnerons celle de Lumel'skii [9] et celle de Newbold et Ho [10].

Lumel'skii a proposé un algorithme de détection de changement brusque des propriétés d'un processus aléatoire. Cet algorithme basé sur des travaux de Shiryaev [14] utilise le critère de Newman-Pearson.

Newbold et Ho présentent un algorithme exact de détection de changement brusque pour les processus à représentation gaussienne markovienne, puis un algorithme continu basé sur le test de Wald. Cet algorithme reprend la même idée que celle utilisée dans Lumel'skii.

Dans ce chapitre, nous présentons un test exact de détection, basé sur le test de Wald pour plusieurs hypothèses. L'expression du rapport de vraisemblance est légèrement différente de celle donnée par Newbold et Ho.

Nous proposons enfin un algorithme simplifié dont nous évaluons les performances.

III. 1. Test exact de détection de changement brusque.

Le test exact exposé dans ce paragraphe est basé sur la méthode séquentielle de Wald pour des hypothèses composées (II. 2. 5.).

III. 1. 1. Présentation du test.

A l'instant $t = o$, le modèle de $y(t)$ est (o).

Ce modèle est susceptible de passer de (o) à (1) à un instant τ.

A l'instant t, nous voulons tester l'hypothèse composée;

$\mathcal{H}^*$: il y a eu changement, par rapport à l'hypothèse ;

H_o : il n'y a pas changement.

Soit un test commençant à $t = t_o$.

A $t = t_n$, nous avons $n + 2$ hypothèses exclusives et exhaustives :

H_o : aucun changement ne s'est produit jusqu'à t_n.

H_{1o} : le changement s'est produit avant le test : $o < \tau < t_o$

H_{1i} : le changement s'est produit pendant le test entre les observations i et i+1, c'est-à-dire : $t_i < \tau < t_{i+1}$.

$\mathcal{H}^*$ est donc composée des hypothèses H_{1i}, (i = o,n) ; de plus :

$$(3\text{-}2) \qquad \sum_{i=o}^{n} p[H_{1i}/\mathcal{H}^*] = 1$$

1°) expression du rapport de vraisemblance généralisé.

Supposons que les hypothèses H_{1i} sachant $\mathcal{H}^*$ soient équiprobables (i = o,n). Ceci n'enlève rien à la généralité du problème.

Compte tenu de (3-2) on a : (3-3) $\quad p[H_{1i}/\mathcal{H}^*] = \frac{1}{n+1} \qquad i = o,n$

Comme $p(H_{1i}/\mathcal{H}^*) > o$, $p(H_{1i}/\mathcal{H}^*)$ (i = o,n) est une fonction $p_1(a)$ définie au chapitre II (II. 2. 5.).

Le rapport de vraisemblance généralisé s'écrit :

$$(3\text{-}4) \qquad \Lambda_n = \frac{1}{n+1} \sum_{i=o}^{n} \frac{p[\mathcal{Y}_n / H_{1i}]}{p[\mathcal{Y}_n / H_o]}$$

2°) Choix des bornes A et B.

Soient : β_i = probabilité de rejeter H_{1i} quand elle est vraie (i = o,n)

β = " " " $\mathcal{H}^*$ " " "

α = " " " H_o " " "

D'après (2-15) et (3-2) : (3-5) $\quad \beta = \frac{1}{n+1} \sum_{i=o}^{n} \beta_i$

et : (3-6)
$$\begin{cases} A = \dfrac{1-\beta}{\alpha} \\ B = \dfrac{\beta}{1-\alpha} \end{cases}$$

III. 1. 2. Calcul du rapport de vraisemblance généralisé.

La présence du signe Σ dans l'expression de Λ_n (3-4) ne permet pas de trouver une relation simple semblable à celle donnée en II. 3. 2. en passant à la forme logarithmique.

Compte tenu de la structure de y(t), nous allons essayer d'exprimer Λ_n de façon récursive.

Il faut tout d'abord déterminer $p[\mathcal{Y}_n / H_o]$ et $p[\mathcal{Y}_n / H_{1i}]$ $i = o,n$

1°) Détermination de $p[\mathcal{Y}_n / H_o]$ et $p[\mathcal{Y}_n / H_{1i}]$ $i = o,n$

On a : (3-7) $p[\mathcal{Y}_n / H] = p[y_n / \mathcal{Y}_{n-1}, H] . p[\mathcal{Y}_{n-1} / H]$

H pouvant être H_o ou H_{1i} , i = o,n.

Nous avons vu (II. 3. 2., a) que l'on peut trouver l'expression de $p[y_n / \mathcal{Y}_{n-1}, H]$ à l'aide du filtre de Kalman correspondant à l'hypothèse H .

Dans le § II. 3. nous avions besoin de deux filtres de Kalman.

Ici, bien qu'il n'y ait que deux structures de base, nous avons besoin d'un filtre par hypothèse, soit n+2 filtres ; chaque filtre comportant un régime transitoire intervenant à un moment différent des autres.

Soient : $\mathcal{J}_o$ le filtre correspondant à l'hypothèse H_o

$\mathcal{J}_{1i}$ " " " H_{1i}, i = o,n

Rappelons que l'hypothèse H_{1i} suppose que le modèle utilisé est (o) jusqu'à l'instant i, puis (1) de l'instant i + 1 à n.

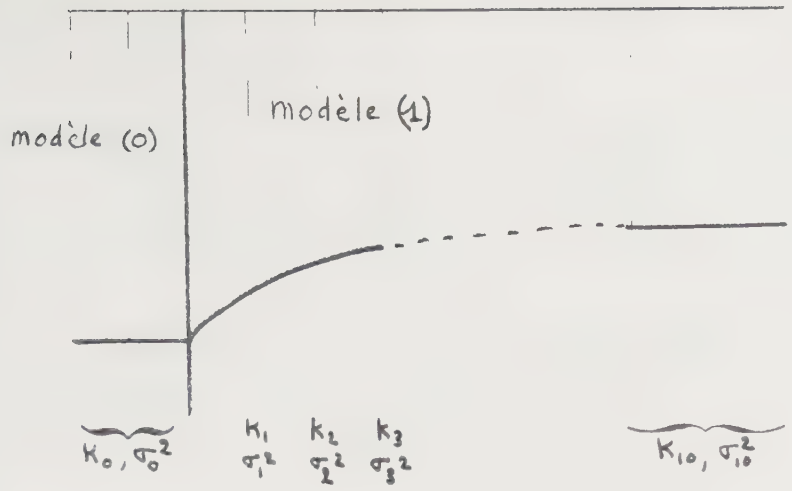
Les relations entre les hypothèses, les modèles et les filtres sont :

(3-8)

hypothèse	filtre	modèle utilisé
H_o	$\mathcal{J}_o$	(o)
H_{1o}	$\mathcal{J}_{1o}$	(1)
H_{11}	$\mathcal{J}_{11}$	(o) jusqu'à i
H_{1i}	$\mathcal{J}_{1i}$	(1) de i+1 à n
H_{1n}	$\mathcal{J}_{1n}$	

Néanmoins, on peut calculer par avance les gains et les variances correspondant au régime transitoire. Ce régime est le même pour tous les filtres, il est seulement décalé dans le temps.

(3-9)



Gains et variances d'erreur corrrespondant au régime transitoire.

Les équations des filtres sont les suivantes :

$\mathcal{J}_0$:

entre deux observations

$$\dot{\hat{X}}_0(t) = F_0 \hat{X}_0(t)$$

à l'instant d'une observation

$$\hat{X}_0(t_{k+}) = \hat{X}_0(t_{k-}) + K_0[y(t_k) - H \hat{X}_0(t_{k-})]$$

$\mathcal{J}_{10}$:

entre deux observations

$$\dot{\hat{X}}_1(t) = F_1 \hat{X}_1(t)$$

à l'instant d'une observation

$$\hat{X}_1(t_{k+}) = \hat{X}_1(t_{k-}) + K_{10}[y(t_k) - H \hat{X}_1(t_{k-})]$$

$\mathcal{J}_{1i}$ $i \geqslant 1$:

mêmes équations que $\mathcal{J}_0$ jusqu'à l'instant i

à partir de l'instant i+1

$$\dot{\hat{X}}_{1i}(t) = F_1 \hat{X}_{1i}(t)$$

$$\hat{X}_{1i}(t_{k+}) = \hat{X}_{1i}(t_{k-}) + K_{k-i}[y(t_k) - H \hat{X}_{1i}(t_{k-})] \quad k > i$$

K_{k-i} est donné par (3-9).

Posons : $\varepsilon_{on} = y(t_n) - H \hat{X}_0(t_{n-})$

$\varepsilon_{1n} = y(t_n) - H \hat{X}_1(t_{n-})$

$\varepsilon_{1i,n} = y(t_n) - H \hat{X}_{1i}(t_{n-})$ et $\sigma_{in} = \begin{cases} \sigma_o & n \leqslant i \\ \sigma_{n-i} & n > i \end{cases}$

Nous obtenons :

$$(3\text{-}10)\quad \begin{cases} p[y(t_n)/\mathcal{Y}_{n-1}, H_o] = \dfrac{1}{\sigma_o\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\frac{\varepsilon_{o\,n}^2}{\sigma_o^2}} \\ p[y(t_n)/\mathcal{Y}_{n-1}, H_{1o}] = \dfrac{1}{\sigma_1\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\frac{\varepsilon_{1\,n}^2}{\sigma_1^2}} \\ p[y(t_n)/\mathcal{Y}_{n-1}/H_{1i}] = \dfrac{1}{\sigma_{n-i}\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\frac{\varepsilon_{1i,n}^2}{\sigma_{n-i}^2}} \quad (i = 1,n) \end{cases}$$

A partir de (3-7) et (3-10), nous pouvons déterminer

$$p[\mathcal{Y}_n/H_o] \quad \text{et} \quad p[\mathcal{Y}_n/H_{1i}]\,(i = o,n)$$

2°) Relation de récurrence du rapport de vraisemblance.

Posons : $\Psi_{n,i} = \dfrac{p[\mathcal{Y}_n/H_{1i}]}{p[\mathcal{Y}_n/H_o]} \quad (i = o,n)$

(3-7) entraine $\Psi_{n,i} = \Psi_{n-1,i} \cdot \dfrac{p[y(t_n)/\mathcal{Y}_{n-1},H_{1i}]}{p[y(t_n)/\mathcal{Y}_{n-1},H_o]}$

La relation de récurrence pour Λ_n est alors :

$$(3\text{-}11)\quad \begin{cases} \Psi_{o,i} = 1 \\ \Psi_{n,i} = \Psi_{n-1,i} \cdot \dfrac{\sigma_o}{\sigma_{n-i}} \cdot e^{-\frac{1}{2}\left[\frac{\varepsilon_{1i,n}^2}{\sigma_{n-i}^2} - \frac{\varepsilon_{on}^2}{\sigma_o^2}\right]} \quad (i = o,n) \\ \Lambda_n = \dfrac{1}{n+1}\sum_{i=o}^{n} \Psi_{n,i} \end{cases}$$

III. 1. 3. Algorithme du test exact.

Les paragraphes précédents nous permettent de déduire un algorithme exact pour la détection d'un changement par saut du modèle de y(t).

Cet algorithme est identique à celui décrit en II. 2. 5. Mais à chaque étape, pour calculer Λ_n donné par (3-11), il est nécessaire de mettre à jour les filtres $\mathcal{J}_o$, $\mathcal{J}_{1i}$, $i = o,n$ afin de calculer les valeurs de ε_{on}, ε_{1n} et $\varepsilon_{1i,n}$ $(i = 1,n)$.

Cet algorithme du test exact, présente plusieurs inconvénients :
- le signe Σ apparaissant dans l'expression du rapport de vraisemblance ne permet pas d'ávoir une forme récursive simple.
- au fur et à mesure que le nombre d'observations augmente, le test exact requiert un nombre de filtres de Kalman croissant. Cela nécessite un volume de calcul de plus en plus grand.

Si la période d'échantillonnage est faible, le test demande un très grand nombre de filtres de Kalman, d'où un grand temps de calcul.
Dans le paragraphe suivant, nous étudions une version simplifiée de ce test, mieux adaptée aux problèmes de détection en temps réel.

III. 2. Algorithme simplifié de détection de changement par saut.

L'une des simplifications du test exact que l'on peut faire est la suivante :
Nous supposons que le changement ne peut se produire pendant le test. :
Deux hypothèses restent à envisager : H_o et H_{1o}.

Le nouveau test a la même forme que celui pour la détection entre deux structures gaussiennes markoviennes. Mais à la différence du test présenté dans le premier chapitre, nous allons utiliser l'idée (dont se servent Newbold et Ho) qui est de faire le meilleur emploi possible de l'information a priori. Ainsi, chaque test commence en utilisant le résultat du test précédent ; la procédure pour la détection de changement étant de faire plusieurs tests, jusqu'à ce que le changement soit détecté.

III. 2. 1. Principe de l'algorithme.

L'information donnée a priori sur le système ou obtenue par les tests précédents est résumée dans la valeur initiale de Λ_n.
Sur cette base, on peut construire deux formes de procédure de détection de changement.

III. 2. 1. 1. Procédure occasionnelle.

Compte tenu de l'hypothèse précédente, le changement est supposé arriver entre les tests. Si le test précédent n'a pas détecté de changement, aux probabilités d'erreur près, alors :

$$p[H_{1o}] = \beta$$

$$p[H_{o}] = 1 - \alpha$$

On choisit comme valeur initiale du nouveau test :

$$\Lambda_o = \lambda(t) \frac{\beta}{1 - \alpha}$$

où $\lambda(t)$, fonction croissante du temps représente la perte d'information entre deux tests.

La procédure est arrêtée dès que le changement est détecté.

Deux inconvénients apparaissent :

- si le changement intervient pendant le test, nous ne savons pas immédiatement ce qui arrive
- s'il intervient entre deux tests, on perd du temps à attendre le prochain test pour le détecter.

Cette procédure est néanmoins utile s'il n'est pas fondamental de détecter rapidement le changement.

III, 2. 1. 2. Procédure continue

On peut définir une procédure continue de la forme précédente. Elle a pour base le règle suivante :

Dès qu'un test se termine par la détection de non-changement, on commence un nouveau test.

La procédure est arrêtée dès qu'il y a détection de changement.

Cette forme de procédure est sensiblement équivalente à la précédente, mais il n'y a pas de perte d'information entre deux tests : $\lambda(t) = 1$

La valeur initiale du nouveau test est donc : $\Lambda_o = \frac{\beta}{1 - \alpha} = B$

Pour le rapport de vraisemblance, nous pouvons alors adopter la règle suivante :

$$\Lambda_n = \max\left\{ \frac{p[y_n/H_{1o}]}{p[y_n/H_o]} , B \right\}$$

La procédure continue ne parait pas a priori cohérente avec l'hypothèse de simplification faite précédemment. Néanmoins elle se justifie lorsqu'elle est considérée comme passage à la limite de la procédure occasionnelle quand la durée entre deux tests tend vers o.
De plus, on montrera par la suite le fait suivant : tant qu'il n'y a pas changement, Λ_n tendra vers B, et quand le changement se produit, Λ_n tend vers A.

Le principe de l'algorithme simplifié que nous proposons est basé sur cette forme de procédure.

III. 2. 2. Description de l'algorithme.

Compte-tenu du principe de l'algorithme, nous faisons plusieurs tests de façon continue.
Pour chaque test, deux hypothèses sont considérées : H_o et H_{1o} . Deux filtres seulement sont nécessaires. Le volume de calcul est donc relativement peu important.
La forme de chaque test est la même que celle décrite en II. 3. 5.

Posons : $\theta_n = \text{Log } \Lambda_n$

$$B = \frac{\beta}{1 - \alpha}$$

L'algorithme simplifié est alors le suivant :
1 - Si nous n'avons aucune information a priori sur les hypothèses, prendre $\theta_o = o$
2 - Suivre les mêmes étapes que celles de l'algorithme décrit en II. 3. 5.
3 - Dès que $\theta_n \leqslant \text{Log B}$, retourner en 2 en prenant $\theta_o = \text{Log B}$.
Dès que $\theta_n \geqslant \text{Log A}$, arrêter la procédure, il y a détection de changement.

Il reste à déterminer les bornes A et B. Mais regardons d'abord la façon dont se comporte l'algorithme.

III. 2. 3. Etude du comportement de l'algorithme.

Pour chaque test, on se base sur la valeur de θ_n variable aléatoire :

$$\theta_n = \text{Log} \frac{p[\mathcal{Y}_n/H_{1o}]}{p[\mathcal{Y}_n/H_o]}$$

A chaque mesure y_n on calcule $p[y_n/\mathcal{Y}_{n-1},H_{1o}]$ et $p[y_n/\mathcal{Y}_{n-1},H_o]$,

puis θ_n : $\theta_n = \theta_{n-1} + \text{Log} \frac{p[y_n/\mathcal{Y}_{n-1},H_o]}{p[y_n/\mathcal{Y}_{n-1},H_{1o}]}$

Posons : $\Delta\theta_n = \text{Log} \frac{p[y_n/\mathcal{Y}_{n-1},H_{1o}]}{p[y_n/\mathcal{Y}_{n-1},H_o]}$

montrons le théorème suivant :

Théorème 3-1 Soient : $E_1[\Delta\theta_n] = E\{\Delta\theta_n/H_{1o}\}$

$E_o[\Delta\theta_n] = E\{\Delta\theta_n/H_o\}$

alors, nous avons les inégalités :

$$(3\text{-}12) \qquad \begin{cases} E_1[\Delta\theta_n] \geqslant o \\ E_o[\Delta\theta_n] \leqslant o \end{cases}$$

Démonstration :

Posons : $p_o = p[y_n/\mathcal{Y}_{n-1},H_o]$

$p_1 = p[y_n/\mathcal{Y}_{n-1},H_{1o}]$ alors $\Delta\theta_n = \text{Log} \frac{p_1}{p_o}$

D'après l'inégalité de Jensen, pour toute fonction $f(x)$ convexe, et pour toute variable aléatoire ξ , on a : $E\,f(\xi) \geqslant f[E(\xi)]$

Prenons la fonction convexe - Log x, on écrit :

$$-E_1 \text{Log} \frac{p_o}{p_1} \geqslant -\text{Log}\,(E_1 \frac{p_o}{p_1})$$

or $E_1 \frac{p_o}{p_1} = \int \frac{p_o}{p_1} p_1 \, d\,y = 1$

donc : $E_1 \text{ Log } \frac{p_0}{p_1} \geq 0$ ou encore : $E_1[\Delta \theta_n] \geq 0$

de la même façon, on montre que : $E_0[\Delta \theta_n] \leq 0$ ∎

Il s'ensuit, compte tenu du principe de l'algorithme que :

θ_n partant de o, va tendre vers l'une des bornes (Log B avec une très forte probabilité : $1 - \alpha$).

Tant que le modèle de y(t) n'a pas changé, on restera proche de Log B, et chaque fois que θ_n atteint Log B, il y a détection de non changement avec une probabilité d'erreur de α.

Dès que le changement intervient, θ_n tendra de Log B vers Log A.

Au début de l'algorithme, il faudra atteindre un certain temps avant que θ_n atteigne Log B (avec une probabilité de $1 - \alpha$).

De même, la détection de changement ne sera pas instantanée. Il y aura un temps de retard entre l'instant τ de changement des paramètres, et l'instant τ' de détection de changement (le temps pour θ_n de passer de Log B à Log A).

θ_n étant aléatoire, ces temps de retard seront aléatoires.

Soient :

- T_b, temps de retard moyen pour atteindre l'un des bornes à partir de l'instant o (début de l'algorithme) sachant qu'il n'y a pas changement.
- T_a, temps de retard moyen de détection de changement.
- T_a et T_b dépendent bien sûr des bornes Log A et Log B choisies, donc des probabilités d'erreur.

Dans le § suivant, nous calculons leur expression.

III. 2. 4. Relations entre les probabilités d'erreur et les temps de retard moyens.

Soient :

- ΔT : période d'échantillonnage de y(t)
- N_a : nombre moyen d'observations nécessaire pour passer de Log B à Log A.
- N_b : " " " " pour atteindre l'une des bornes à partir de l'instant o sachant qu'il n'y a pas changement.

Nous avons les relations :

$$(3\text{-}13)\quad \begin{cases} T_a = N_a \cdot \Delta T \\ T_b = N_b \cdot \Delta T \end{cases}$$

Pour calculer N_a et N_b, nous avons besoin de calculer $E[\theta_n/H_o]$ et $E[\theta_n/H_{1o}]$.

1°) Calcul de $E[\theta_n/H_o]$ et $E[\theta_n/H_{1o}]$

Nous faisons l'hypothèse que la durée du régime transitoire intervenant dans les équations du filtre de Kalman est négligeable devant l'ordre de grandeur des nombres N_a et N_b. Les filtres $\mathcal{J}_o$ et $\mathcal{J}_{1o}$ sont donc supposés en régime stationnaire.

Théorème 3.2

$E[\theta_n/H_o]$ et $E[\theta_n/H_{1o}]$ sont des fonctions linéaires de n.

Démonstration

Calculons $E[\theta_n/H_o]$:

$$\theta_n = \theta_{n-1} + \text{Log}\,\frac{\sigma_o}{\sigma_1} - \frac{1}{2}\left[\frac{\varepsilon_{1n}^2}{\sigma_1^2} - \frac{\varepsilon_{on}^2}{\sigma_o^2}\right]$$

d'où il vient :

$$E[\theta_n/H_o] = E[\theta_{n-1}/H_o] + \text{Log}\,\frac{\sigma_o}{\sigma_1} - \frac{1}{2}\left[\frac{E[\varepsilon_{1n}^2/H_o]}{\sigma_1^2} - \frac{E[\varepsilon_{on}^2/H_o]}{\sigma_o^2}\right]$$

Sous l'hypothèse H_o, y(t) a pour modèle (o)

On utilise le filtre $\mathcal{J}_o$ pour calculer ε_{on} ; c'est le filtre optimal . L'innovation ε_{on} est un bruit blanc gaussien (Kailath [7]).

On obtient :

$$E[\varepsilon_{on}^2/H_o] = \sigma_o^2$$

Pour calculer ε_{1n}, on utilise le filtre $\mathcal{J}_{1o}$ (basé sur le modèle 1) ; ce n'est pas le filtre optimal.

On a : $\varepsilon_{1n} = H\,(X(t_n) - \hat{X}_1(t_n)) + w(t_n)$

Soit : $E(\varepsilon_{1n}^2) = H\,E\{(X(t_n) - \hat{X}_1(t_n))(X(t_n) - \hat{X}_1(t_n))'\}\,H' + R$

Posons : $\Sigma_{o1}(t_n) = E\{(X(t_n) - \hat{X}_1(t_n))\cdot(X(t_n) - \hat{X}_1(t_n))'\}$

Le calcul de $\Sigma_{o1}(t_n)$ connaissant les modèles (o) et (1) est exposé en appendice [III]. Compte tenu de l'hypothèse faite précèdemment, nous ne considérons pour $\Sigma_{o1}(t_n)$ que la valeur d'équilibre $\Sigma_{o1}(\infty) = \Sigma_{o1}$.

Σ_{o1} est indépendant de n (voir appendice [III]).

On peut maintenant écrire :

$$E[\theta_n/H_o] = E[\theta_{n-1}/H_o] + \text{Log}\,\frac{\sigma_o}{\sigma_1} - \frac{1}{2}\left[\frac{H\Sigma_{o1}H'+R}{\sigma_1^2} - 1\right]$$

ou :

$$(3\text{-}14) \qquad E[\theta_n/H_o] = n\left[\text{Log}\,\frac{\sigma_o}{\sigma_1} - \frac{1}{2}\left[\frac{H\Sigma_{o1}H'+R}{\sigma_1^2} - 1\right]\right]$$

$E[\theta_n/H_o]$ est donc une fonction linéaire de n.

De la même façon, on montre que $E[\theta_n/H_{1o}]$ est une fonction linéaire de n :

Posons : $\Sigma_{1o} = E\{(X(t_n) - \hat{X}_o(t_n))\cdot(X(t_n) - \hat{X}_o(t_n))'\}$

$E[\theta_n/H_{1o}]$ s'écrit :

$$(3\text{-}15) \qquad E[\theta_n/H_{1o}] = n\left[\text{Log}\,\frac{\sigma_o}{\sigma_1} - \frac{1}{2}\left[1 - \frac{H\Sigma_{1o}H'+R}{\sigma_o^2}\right]\right]$$

Posons :

$$(3\text{-}15\text{bis}) \begin{cases} E_o(\theta) = \text{Log}\,\dfrac{\sigma_o}{\sigma_1} - \dfrac{1}{2}\left[\dfrac{H\Sigma_{o1}H'+R}{\sigma_1^2} - 1\right] \\ E_1(\theta) = \text{Log}\,\dfrac{\sigma_o}{\sigma_1} - \dfrac{1}{2}\left[1 - \dfrac{H\Sigma_{1o}H'+R}{\sigma_o^2}\right] \end{cases}$$

2°) Calcul de N_a et N_b.

Soit N le point d'arrêt du test, N est une variable aléatoire. Pour atteindre l'une des bornes (au début de l'algorithme), $E[\theta_N/H_o]$ est telle que :

$$E[\theta_N/H_o] = \alpha \text{ Log } A + (1-\alpha) \text{ Log } B$$

Posons : $N_b = E[N/H_o]$

Pour n fixe on a : $E[\theta_n/H_o] = n\, E_o(\theta)$

D'où : $E[\theta_N/H_o] = N_b\, E_o(\theta)$

et,

$$(3\text{-}16) \qquad N_b = \frac{\alpha \text{ Log } A + (1-\alpha)\text{Log } B}{E_o(\theta)}$$

De même, pour que l'algorithme détecte un changement, il faut que θ_n passe de Log B à Log A.

Soit en prenant les espérances :

$$E[\theta_N/H_1] = \text{Log } A - \text{Log } B$$

D'où par un raisonnement similaire à N_B, on obtient :

$$(3\text{-}17) \qquad N_a = \frac{\text{Log } A - \text{Log } B}{E_1(\theta)}$$

3°) Relations entre T_a, T_b et les probabilités d'erreur α, β.

Les bornes A et B sont liées aux probabilités d'erreur α, β par les équations :

$$(3\text{-}18) \qquad A = \frac{1-\beta}{\alpha} \quad , \quad B = \frac{\beta}{1-\alpha}$$

De (3-16), (3-17) et (3-18), on déduit :

$$(3\text{-}19) \qquad \begin{cases} T_b = \dfrac{\alpha \text{Log} \dfrac{1-\beta}{\alpha} + (1-\alpha)\text{Log}\dfrac{\beta}{1-\alpha}}{E_o(\theta)} \cdot \Delta T \\ T_a = \dfrac{\text{Log}\dfrac{1-\beta}{\alpha} - \text{Log}\dfrac{\beta}{1-\alpha}}{E_1(\theta)} \cdot \Delta T \end{cases}$$

III. 2. 5. Choix des bornes A, B.

Ce qui importe dans l'algorithme est d'une part la rapidité de la détection de changement, (c'est-à-dire T_a), d'autre part l'exactitude de la détection, (c'est-à-dire les probabilités d'erreurs).
Les deux tendances sont bien sur contradictoires, mais quelque soit le but recherché, les relations (3-18) et (3-19) nous permettrons de déterminer les bornes A, B.

III. 3. Conclusion

Le test exact présenté nécessite un volume et un temps de calcul croissants avec le nombre d'observations.
L'algorithme simplifié proposé est pratiquement identique en temps et volume de calcul à l'algorithme de détection entre deux structures gaussiennes markoviennes.
Cet algorithme de détection de changement par saut est donc mieux adapté au temps réel.

Newbold et Ho [10] présente une méthode de calcul des performances de l'algorithme (nombre moyen d'observations nécessaire pour détecter) pour le cas où le paramètre susceptible de varier est Q avec Q scalaire.
Nous avons traité dans ce chapitre le cas plus général où plusieurs paramètres peuvent brusquement changer. La méthode de calcul du nombre moyen d'observation est relativement simple.
Enfin, nous avons mis l'accent sur le choix à faire entre la rapidité de détection de l'instant τ de changement et l'exactitude de la détection. Ce choix intervient dans la détermination des bornes A et B.

L'algorithme simplifié que nous venons de présenter suppose connue la valeur finale des paramètres après changement. Mais en pratique, on ignore généralement la valeur exacte après changement. On sait seulement qu'ils appartiennent à un certain ensemble de valeurs. Les hypothèses mises en jeu ne sont plus "simples" ; ce sont des hypothèses "multiples".
Avant d'aborder ce problème, nous allons d'abord étudier le comportement du test entre deux structures gaussiennes markoviennes quand le vrai modèle de y(t) n'est ni (o), ni (1).

Chapitre IV

Etude du comportement du test de détection entre deux structures gaussiennes-markoviennes dans le cas général

Nous avons présenté dans le chapitre II un test de détection entre deux structures gaussiennes markoviennes.

Dans le troisième chapitre, pour traiter le problème de changement par saut, nous avons utilisé le test de détection entre deux structures gaussiennes markoviennes.

Ce test supposait exacte la connaissance des paramètres des deux modèles envisagés : Le vrai modèle de $y(t)$ était (o) ou (1). Mais dans la plupart des cas, en pratique, le vrai modèle suivi par $y(t)$ ne correspond à aucun des deux modèles escomptés. Il est alors utile de connaitre la sensibilité du test de détection aux erreurs de modèles, en particulier, de savoir comment il réagit lorsque le vrai modèle de $y(t)$ ne correspond ni à l'un ni à l'autre des modèles envisagés.

Dans ce chapitre, nous étudions le comportement du test de détection quand le vrai modèle du signal observé $y(t)$ n'est ni (o) ni (1) ; en particulier :

- le nombre moyen d'observations nécessaire pour prendre une décision,
- la probabilité d'erreur de détection.

Nous nous limiterons dans cette étude à une certaine forme de modèle. Par exemple, si a est le paramètre susceptible de changer, et si :

a_o correspond au modèle (o)
a_1 " " (1)
a " au vrai modèle de $y(t)$,

alors nous imposerons que a soit lié à a_o et a_1 par une relation du type :

$$a = (1-\zeta)\, a_o + \zeta\, a_1 \quad , \quad \zeta \in R \tag{4-1}$$

on introduit une relation d'ordre sur a qui est la suivante :

$a_\alpha > a_\beta$ si et seulement si $\zeta_\alpha > \zeta_\beta$.

Soient : Ω, l'ensemble des valeurs possibles de a (vérifiant (4-1))

$\Omega(a_o, a_1)$, le sous-ensemble de Ω tel que si $a \in \Omega(a_o, a_1)$ les conditions de stabilité ou de positivité nécessaires pour a soient satisfaites.

Nous supposerons toujours dans la suite de cette étude que $a \in \Omega(a_o, a_1)$.

Par exemple si a = F, F_o et F_1 asymptotiquement stables, alors $\Omega(F_o, F_1)$ sera l'ensemble des valeurs de F telles que $F = \zeta F_1 + (1-\zeta)F_o$ soit asymptotiquement stable.

Soit $\omega(a_o, a_1)$ l'ensemble des valeurs de ζ telles que :

si $\zeta \in \omega(a_o, a_1)$, alors $a(\zeta) \in \Omega(a_o, a_1)$.

Les propriétés de l'innovation dans le cas d'un filtre sous-optimal nous permettront d'étudier les statistiques du rapport de vraisemblance. Nous montrerons alors la convergence du test de détection dans le cas général.

Puis nous donnerons une expression de la valeur moyenne du nombre d'échantillons nécessaire pour prendre une décision.

Enfin, nous proposerons une règle de troncature du test pour le cas où le nombre d'observations devient trop grand.

IV. 1. L'innovation dans le cas d'un filtre sous-optimal.

IV. 1. 1. Modèle représentant l'innovation ou pseudo-innovation.

Le modèle vrai représentant le signal scalaire observé y(t) est :

$$(4\text{-}2)\qquad \begin{aligned} \dot{X}(t) &= F\,X(t) + v(t) \qquad & \text{avec } EX(t_o)X'(t_o) &= P \\ y(t_n) &= HX(t_n) + w(t_n) & E\,v\cdot v' &= Q\,\delta \\ & & Ew_i w_j &= R\delta ij \text{ (observation discrète)} \end{aligned}$$

Le modèle utilisé servant à définir l'algorithme de filtrage est :

$$(4\text{-}2 \text{ bis})\qquad \begin{aligned} \dot{X}_f(t) &= F_f X_f(t) + v(t) \qquad & \text{avec } EX_f(t_o)X'_f(t_o) &= P_f \\ y(t_n) &= H\,X_f(t_n) + w(t_n) & Ev_f v'_f &= Q_f\,\delta \end{aligned}$$

Soit le processus :

$$(4\text{-}3) \qquad \varepsilon_n = y(t_n) - H\,\hat{X}_f(t_{n-})$$

où $\hat{X}_f(t_{n-}) = E\{X(t_n)/y(t_o)\cdots,y(t_{n-1}),\ F_f,\ Q_f\}$.

Nous allons chercher le modèle représentant ε_n.

Pour y(t), nous avons considéré un modèle continu à observations discrètes, mais pour la commodité des démonstrations qui vont suivre, nous utiliserons un modèle discret. Cela n'enlève rien à la généralité des résultats. En effet, il existe toujours un modèle discret qui représente exactement le modèle continu.
Les équations des deux modèles discrets représentant (4-2) et (4-2 bis) sont :

$$(4\text{-}4) \qquad \begin{cases} X(t_{n+1}) = \Phi\, X(t_n) + v_d(t_n) \\ y(t_n) = H\, X(t_n) + w(t_n) \end{cases} \qquad \text{avec} \quad E\{v_d(t_n)\, v'_d(t_m)\} = Q_d\, \delta_{n,m}$$

$$(4\text{-}5) \qquad \begin{cases} \hat{X}_f(t_{n+1}) = \Phi_f\, \hat{X}_f(t_n) + v_{df}(t_n) \\ y(t_n) = H\, \hat{X}_f(t_n) + w(t_n) \end{cases} \qquad \text{avec} \quad E\{v_{df}(t_n)\, v'_{df}(t_m)\} = Q_{df}\, \delta_{n,m}$$

Les relations avec le modèle continu sont :

$$\Phi = e^{F\Delta t}$$

$$\Phi_f = e^{F_f\Delta t}$$

$$Q_d = \int e^{F(\Delta t - s)}\, Q\, e^{F'(\Delta t - s)}\, d\, s$$

$$\Delta t = t_{n+1} - t_n$$

Les équations de l'algorithme de filtrage utilisant (4-5) sont :
(appendice II, 2è partie).

$$(4\text{-}6) \qquad \hat{X}_f(t_{n+1-}) = \Phi_f\, \hat{X}_f(t_{n-}) + \Phi_f\, K_f[y(t_n) - H\, \hat{X}_f(t_{n-})]$$

Nous supposons comme dans le chapitre précédent que le filtre a atteint son régime stationnaire.

Compte tenu de (4-4), (4-6) et (4-3), ε_n est donné par :

$$(4\text{-}7)\quad \begin{cases} \begin{bmatrix} X(t_{n+1}) \\ \hat{X}_f(t_{n+1-}) \end{bmatrix} = \begin{bmatrix} \Phi & 0 \\ \Phi_f K_f H & \Phi_f(I-K_f H) \end{bmatrix} \begin{bmatrix} X(t_n) \\ \hat{X}_f(t_{n-}) \end{bmatrix} + \begin{bmatrix} v_d(t_n) \\ \Phi_f K_f w(t_n) \end{bmatrix} \\ \varepsilon_n = [H, -H] \begin{bmatrix} X(t_n) \\ X_f(t_{n-}) \end{bmatrix} + w(t_n). \end{cases}$$

ou encore, en posant :

$$X_n = \begin{bmatrix} X(t_n) \\ X_f(t_{n-}) \end{bmatrix}, \quad \Phi = \begin{bmatrix} \Phi & 0 \\ \Phi_f K_f H & \Phi_f(I-K_f H) \end{bmatrix}, \quad H = [H, -H]$$

et $\gamma_n = \begin{bmatrix} v_d(t_n) \\ \Phi_f K_f w(t_n) \end{bmatrix}$

on obtient :

$$(4\text{-}7\text{bis})\quad \begin{cases} X_{n+1} = \Phi X_n + \gamma_n \\ \varepsilon_n = H X_n + w(t_n) \end{cases}$$

On en déduit les propriétés de ε_n :

Théorème 4.1.

ε_n est un processus à représentation gaussienne markovienne.

Démonstration.

Compte tenu des relations (4-7bis), il suffit de montrer que Φ est asymptotiquement stable.

Φ est une matrice triangulaire inférieure par blocs.

L'ensemble des valeurs propres de Φ est la concaténation des valeurs propres de Φ et des valeurs propres de $\Phi_f(I-K_f H)$, (voir appendice IV).

Or Φ et Φ_f sont asymptotiquement stables, $\Phi_f(I-K_fH)$ l'est également (FAURRE [5]). Φ est donc asymptotiquement stable.

Cette propriété, compte tenu de (4-7bis) suffit à démontrer que le processus ε_n est un processus à représentation gaussienne markovienne.

On en conclut à propos du processus ε_n :

- si les modèles donnés par (4-4) et (4-5) sont les mêmes, alors le filtre donné par l'équation (4-6) est optimal.
ε_n est un bruit blanc gaussien (Kailath [7]).

- si les modèles (4-4) et (4-5) différent par la valeur d'un ou plusieurs paramètres, ε_n n'est plus une innovation, on l'appelle pseudo innovation, le filtre donné par (4-6) étant sous-optimal.
ε_n est un processus à représentation gaussienne-markovienne.

IV. 1. 2. Propriétés de la fonction d'autocorrélation de la pseudo-innovation.

Soit $M = E\{X_o X'_o\}$

Compte tenu de (4-7bis), M est donné par l'équation de Lyapunov :

$$M - \Phi M \Phi' = \mathcal{Q} \quad \text{où} \quad \mathcal{Q} = \begin{bmatrix} Q_d & 0 \\ 0 & \Phi_f K_f R K'_f \Phi'_f \end{bmatrix}$$

Toujours d'après (4-7bis):

$$X_n = X_{n-k} + \sum_{j=1}^{k} \Phi^{j-1} V_{n-j}$$

Alors :

$$E\{X_n X'_{n-k}\} = \Phi^k M \quad \text{et de même :} \quad E\{X_n \cdot w_{n-k}\} = \Phi^{k-1} \begin{bmatrix} 0 \\ \Phi_f K_f R \end{bmatrix}$$

La fonction d'autocorrélation est donnée par :

$$(4\text{-}8) \quad \begin{cases} E[\varepsilon_n \cdot \varepsilon_{n-k}] = H \Phi^k M H' + H \Phi^{k-1} \mathcal{R} & k > o \\ E(\varepsilon_n^2) = H M H' + R \end{cases}$$

où $\mathcal{R} = \begin{bmatrix} 0 \\ \Phi_f K_f R \end{bmatrix}$

Posons :

$$C_k = E[\varepsilon_n . \varepsilon_{n-k}]$$
$$C_o = E[\varepsilon_n^2]$$

Remarquons que : $C_k = C_{-k}$

La fonction d'autocorrélation est symétrique.

Une autre propriété immédiate est donnée par le théorème suivant :

Théorème 4.2 La fonction d'autocorrélation de ε_n est de carré sommable.

ou encore : $\sum_{k=o}^{\infty} C_k^2 < +\infty$

Démonstration : $C_o = H M H' + R$

$$C_k = H \Phi^{k-1} [\Phi M H' + \mathcal{R}] \qquad k > o$$

Posons : $L = \Phi M H' + \mathcal{R}$

$B = L L'$ B est symétrique non négative.

alors $C_k^2 = H \Phi^{k-1} B \Phi^{k-1'} H'$

et $\sum_{k=o}^{\infty} C_k^2 = C_o^2 + H [\sum_{k=1}^{\infty} \Phi^{k-1} B \Phi^{k-1'}] H'$

Φ est asymptotiquement stable (démonstration théorème 4-1).

alors, $T = \sum_{k=1}^{\infty} \Phi^{k-1} B \Phi^{k-1'}$ est solution de l'équation de Lyapunov :

$$T - \Phi T \Phi' = B \qquad \text{et} \quad T > o$$

et $\sum_{k=o}^{\infty} C_k^2 = C_o^2 + H T H'$

La suite C_k est de carré sommable. ■

IV. 1. 3. Estimation de $E[\varepsilon_n^2]$ à partir de la réalisation de ε_n.

Nous présentons dans ce paragraphe des résultats qui nous serviront pour la suite de l'étude.

Soit (4-9) $\hat{C}_{on} = \frac{1}{n} \sum_{i=1}^{n} \varepsilon_i^2$

C_{on} est une variable aléatoire non gaussienne, Calculons ses deux premiers moments.

On a :

$$(4\text{-}10) \qquad E\ \hat{C}_{o,n} = C_o$$

d'après Bartlett [1] :

$$\text{var}\ \hat{C}_{on} = \frac{1}{n} \sum_{k=-n+1}^{\lambda-1} \left(1 - \frac{k}{n}\right)\ \mathcal{E}^{\circ}_{(k)}$$

$$\mathcal{E}^{\circ}_{(k)} = 2\, C_k^2 + K_k$$

K_k est le cumulant d'ordre 4 de ε_n

Or nous avons les propriétés suivantes :

1 - $C_k = C_{-k}$

2 - $K_f = o$ Le cumulant d'ordre 4 d'un processus gaussien est nul.

(Parzen [11]).

Alors il vient :

$$(4\text{-}11) \qquad \text{var}\ \hat{C}_{on} = \frac{4}{n} \sum_{k=o}^{n-1} \left(\frac{n-k}{n}\right)\ C_k^2$$

D'après le théorème 4-2, C_k est de carré sommable.

en posant $\gamma = 4 \sum_{k=o}^{\infty} C_k^2$

Alors :

$$(4\text{-}12) \qquad \text{var}\ \hat{C}_{on} \leqslant \frac{\gamma}{n}$$

Dans ce paragraphe, les démonstrations ont été faites en supposant que seul le paramètre F différait dans les deux modèles. Cela n'enlève rien à la généralité des résultats.

IV. 2. Statistiques du rapport de vraisemblance.

Les propriétés de la pseudo innovation dans le cas d'un filtre sous optimal vont nous permettre de déduire des résultats sur les statistiques du rapport de vraisemblance θ_n.

Nous posons (comme au chapitre II) :

$$\theta_n = \text{Log} \frac{p[y_n/H_1]}{p[y_n/H_o]}$$

Les deux filtres utilisés sont $\mathcal{J}_o$ et $\mathcal{J}_1$ correspondant aux modèles (o) et (1). Nous faisons, pareillement au chapitre III, l'hypothèse que la durée des régimes transitoires est négligeable.

Soient ε_{on} et ε_{1n}, les pseudo innovations correspondant à ces filtres.

D'après (2-20) :

$$\theta_n = \theta_{n-1} + \text{Log} \frac{\sigma_o}{\sigma_1} - \frac{1}{2}\left[\frac{\varepsilon_{1n}^2}{\sigma_1^2} - \frac{\varepsilon_{on}^2}{\sigma_o^2} \right]$$

Nous nous limitons aux modèles de la forme :

$$(4\text{-}1) \quad a(\zeta) = \zeta a_1 + (1-\zeta)a_o, \quad \zeta \in \omega(a_o, a_1)$$

Soient :

$$E_\zeta \theta_n = E[\theta_n/\zeta]$$

$$\text{var}_\zeta \theta_n = E[(\theta_n - E_\zeta \theta_n)^2/\zeta]$$

IV. 2. 1. Propriétés de $E_\zeta \theta_n$.

Une première propriété de $E_\zeta \theta_n$ est donnée par le théorème suivant :

Théorème 4-3

$E_\zeta \theta_n$ est une fonction linéaire de n.

Démonstration

θ_n peut s'écrire sous la forme :

$$(4\text{-}13) \qquad \theta_n = n \text{ Log} \frac{\sigma_o}{\sigma_1} - \frac{n}{2}\left[\frac{1}{n} \frac{1}{\sigma_1^2} \sum_{i=1}^{n} \varepsilon_{1i}^2 - \frac{1}{n} \frac{1}{\sigma_o^2} \sum_{i=1}^{n} \varepsilon_{oi}^2 \right]$$

ou encore :

$$(4\text{-}14) \qquad \theta_n = n \operatorname{Log} \frac{\sigma_o}{\sigma_1} - \frac{n}{2} \left[\frac{\hat{C}_{on,1}}{\sigma_1^2} - \frac{\hat{C}_{o,n,o}}{\sigma_o^2} \right]$$

De (4-10) il vient :

$$E_\zeta \, \theta_n = n \operatorname{Log} \frac{\sigma_o}{\sigma_1} - \frac{n}{2} \left[\frac{C_{o1}}{\sigma_1^2} - \frac{C_{oo}}{\sigma_o^2} \right]$$

soit :

$$(4\text{-}15) \qquad E_\zeta \, \theta_n = n \left[\operatorname{Log} \frac{\sigma_o}{\sigma_1} - \frac{1}{2} \left[\frac{C_{o1}}{\sigma_1^2} - \frac{C_{oo}}{\sigma_o^2} \right] \right]$$

avec : $C_{o1} = E \, \varepsilon_{1n}^2$

$C_{oo} = E \, \varepsilon_{on}^2$

L'expression de C_{oo} et C_{o1} a été explicitée en (4-8), le calcul est exposé au chapitre VI.

Posons :

$$(4\text{-}16) \qquad E_\zeta \, \theta = \operatorname{Log} \frac{\sigma_o}{\sigma_1} - \frac{1}{2} \left[\frac{C_{o1}}{\sigma_1^2} - \frac{C_{oo}}{\sigma_o^2} \right]$$

alors :

$$(4\text{-}17) \qquad E_\zeta \, \theta_n = n \; E_\zeta \, \theta$$

Une autre propriété sur $E_\zeta \, \theta$, donc sur $E_\zeta \, \theta_n$ est la suivante :

Théorème 4-4

L'application $\zeta \rightarrow E_\zeta \, \theta$ est monotone croissante.

Démonstration.

Nous faisons intervenir le paramètre susceptible de varier, a.

Si a = F, nous avons calculé l'expression de $E_\zeta \, \theta$ pour F scalaire, puis démontré le propriété de monotonie croissante (voir appendice [V]).

si a = Q

Les deux modèles (o) et (1) ne différent que par la covariance du bruit d'excitation.

Soient : Q_i, correspondant au modèle i, $(i = o,1)$

Q, " au vrai modèle de y(t) Q est telle que :

$$Q = \zeta Q_1 + (1-\zeta) Q_o$$

$\omega(Q_o, Q_1)$, l'ensemble des valeurs de ζ telles que $Q \geq o$

on suppose $\zeta \in \omega(Q_o, Q_1)$

Soient encore : Σ_i, la covariance vraie d'erreur de filtrage quand $Q = Q$ et le filtre utilisé est $\mathcal{J}_i$

Σ_{iu}, " d'erreur de filtrage quand $Q = Q_i$ et le filtre utilisé est $\mathcal{J}_i$.

$$\delta\Sigma_i = \Sigma_i - \Sigma_{iu} \quad (i = o,1)$$

alors on a : $\sigma_i^2 = H \Sigma_{iu} H' + R \quad (i = o,1)$

$E_\zeta \theta$ s'écrit : $$E_\zeta \theta = \text{Log} \frac{\sigma_o}{\sigma_1} - \frac{1}{2}\left[\frac{H\Sigma_1 H' + R}{\sigma_1^2} - \frac{H\Sigma_o H' + R}{\sigma_o^2} \right]$$

ou encore : $$E_\zeta \theta = \text{Log} \frac{\sigma_o}{\sigma_1} - \frac{1}{2}\left[\frac{H\delta\Sigma_1 H'}{\sigma_1^2} - \frac{H\delta\Sigma_o H'}{\sigma_o^2} \right] \quad (4\text{-}18)$$

Σ_i et Σ_{iu} verifient des équations de Riccati discrètes, les observations étant discrètes. Mais toute équation de Riccati discrète peut être approchée au premier ordre en Δt par une équation différente de Riccati (Faurre [4]).

alors,

Σ_i et Σ_{iu} vérifient les équations :

$$\left.\begin{aligned} (F-K_iH)\Sigma_i + \Sigma_i(F-K_iH)' + K_iRK_i' + Q &= o \\ (F-K_iH)\Sigma_{iu} + \Sigma_{iu}(F-K_iH)' + K_iRK_i' + Q_i &= o \end{aligned}\right\} (i=o,1)$$

soit pour $\delta\Sigma_i$:

$$(F-K_iH)\delta\Sigma_i + \delta\Sigma_i(F-K_iH)' + a_i(Q_1-Q_o) = o$$

ou $a_1 = \zeta - 1$

$a_o = \zeta$

posons $\delta'\Sigma_i \, a_i = \delta\Sigma_i$

alors $\delta'\Sigma_i$ vérifie l'équation : $(F-K_iH)\delta'\Sigma_i + \delta'\Sigma_i(F-K_iH)' + (Q_1-Q_o) = o \quad (i = o,1)$

$\delta'\Sigma_i$ est donc indépendant de ζ.

$\delta\Sigma_1$ et $\delta\Sigma_o$ sont des fonctions linéaires de ζ.

de (4-18) on peut écrire :

$$E_\zeta\ \theta = \text{Log}\ \frac{\sigma_o}{\sigma_1} - \frac{1}{2}\left[\zeta\left(\frac{H\delta'\Sigma_1 H'}{\sigma_1^2} - \frac{H\delta'\Sigma_o H'}{\sigma_o^2}\right) - \frac{H\delta'\Sigma_1 H'}{\sigma_1}\right]$$

$E_\zeta\ \theta$ est une fonction monotone de ζ.

De plus, on a montré (théorème 3-1) $\quad \zeta = 1 \quad E_\zeta\ \theta = E_1\ \theta \geqslant o$

$$\zeta = o \quad E_\zeta\ \theta = E_o\ \theta \leqslant o$$

$E_\zeta\ \theta$ est donc une fonction monotone croissante de ζ.

■

IV. 2. 2. Propriétés de var $_\zeta\ \theta_n$

Théorème 4-5 : var $_\zeta \theta_n$ est majorée par une fonction linéaire de n.

Démonstration :

$$\text{var}_\zeta\ \theta_n = E(\theta_n - E_\zeta\ \theta_n)^2$$

Compte tenu de (4-14) et (4-15)

$$\text{var}_\zeta\ \theta_n = \frac{n^2}{4} E\left(\frac{\hat{C}_{o,n,1} - C_{o1}}{\sigma_1^2} - \frac{\hat{C}_{o,n,o} - C_{oo}}{\sigma_o^2}\right)^2$$

$$= \frac{n^2}{4}\left[\frac{1}{\sigma_1^4}\text{var}\ \hat{C}_{on,1} + \frac{1}{\sigma_o^4}\text{var}\ \hat{C}_{on,o} - \frac{2}{\sigma_o^2\sigma_1^2} E(\hat{C}_{on,1} - C_{o1})(\hat{C}_{on,o} - C_{oo})\right]$$

or $|EXY| < \sqrt{\text{varX. var Y}}$

ce qui implique : $\frac{n^2}{4}\left[\frac{\sqrt{\text{var}\hat{C}_{on,1}}}{\sigma_1^2} - \frac{\sqrt{\text{var}\hat{C}_{on,o}}}{\sigma_o^2}\right]^2 \leqslant \text{var}_\zeta\ \theta \leqslant \frac{n^2}{4}\left[\frac{\sqrt{\text{var}C_{on,1}}}{\sigma_1^2} - \frac{\sqrt{\text{var}C_{on,o}}}{\sigma_o^2}\right]^2$

calculons l'expression de var $\hat{C}_{on,1}$ et var $\hat{C}_{on,o}$:

D'après (4-12) : var $\hat{C}_{on,i} \leqslant \frac{\gamma_i}{n}$ où $\gamma_i = 4 \sum_{k=o}^{\infty} C_{k_i}^2$

$$C_{k_i} = E(\varepsilon_{i,n} \cdot \varepsilon_{i,n-k}) \quad i = o,1$$

De plus, nous savons (théorème 4-2) que $\gamma_i < +\infty$

on obtient : $$\text{var}_\zeta \theta_n < \frac{n}{4}\left(\frac{\sqrt{\gamma_1}}{\sigma_1^2} + \frac{\sqrt{\gamma_o}}{\sigma_o^2} \right)^2 \qquad (4\text{-}19)$$

IV. 3. Convergence du test dans le cas général.

Nous montrons dans ce paragraphe que le nombre d'observations nécessaires pour détecter est toujours fini.

Théorème 4-6

Le test de détection entre deux structures gaussiennes markoviennes, dans le cas général, se termine presque surement en un temps fini.

Démonstration :

Nous supposons pour la généralité du problème que le vrai modèle n'est ni (o), ni (1).

Nous avons montré (théorèmes 4-3 et 4-5) : $E_\zeta \theta_n = n \cdot E_\zeta \theta$

$$\text{var}_\zeta \theta_n \leqslant n\gamma$$

où $$\nu = \frac{1}{4}\left(\frac{\sqrt{\gamma_1}}{\sigma_1^2} + \frac{\sqrt{\gamma_o}}{\sigma_o^2} \right)^2$$

Nous savons également (théorème 2-2) que $E_\zeta\theta$ est une fonction monotone de ζ suivant la valeur de ζ, on aura $E_\zeta\theta > o$ ou $E_\zeta\theta < o$.

Montrons un premier résultat :

Si $E_\zeta\theta > o$

Soit N_1 tel que $n > N_1 \rightarrow E_\zeta \theta_n > \text{Log } A$

alors pour $n > N_1$: $p[\theta_n < \text{Log } A] = p[\theta_n - E_\zeta\theta_n < \text{Log } A - E_\zeta\theta_n]$

$$\leqslant p[|\theta_n - E_\zeta\theta_n| > E_\zeta\theta_n - \text{Log } A]$$

$$\leqslant \frac{\text{var}_\zeta \theta_n}{(E_\zeta\theta_n - \text{Log}A)^2} \qquad \text{d'après Bienaymé-Tchebitchev}$$

Comme : $\text{var}_\zeta \theta_n \leqslant n\nu$

et $(E_\zeta\theta_n - \text{Log}A)^2 = (nE_\zeta\theta - \text{Log}A)^2$

quand $n \to \infty$, $(E_\zeta \theta_n - \text{Log } A)^2 \sim n^2 E_\zeta \theta^2$

ce qui implique : $p[\theta_n < \text{Log } A] \leqslant \frac{1}{n} \frac{\nu}{E_\zeta \theta^2}$

ou $\lim_{n \to \infty} p[\theta_n < \text{Log } A] = o$ (4-20)

Si $E_\zeta < o$

Soit N_o tel que $n > N_o \to E_\zeta\ \theta_n > \text{Log } B$

pour $n > N_o$: $p[\theta_n > \text{Log } B] = p[\theta_n - E_\zeta\ \theta_n > \text{Log } B - E_\zeta\ \theta_n]$

$\leqslant \frac{\text{var}_\zeta\ \theta_n}{(\text{Log}B - E_\zeta \theta_n)^2}$ d'après Bienaymé - Tchebitchev

quand $n \to \infty$ $(\text{Log } B - E_\zeta\ \theta_n)^2 \sim n^2\ E_\zeta\ \theta^2$

ce qui entraine :

$$p[\theta_n > \text{Log } B] \leqslant \frac{1}{n}\ \frac{\nu}{E_\zeta \theta^2}$$

ou encore : $\lim_{n \to \infty} p[\theta_n > \text{Log } B] = o$ (4-21)

Soit N le point d'arrêt du test (abscisse aléatoire pour laquelle θ_n a franchi pour la première fois l'une des deux bornes).

Supposons $E_\zeta\ \theta > o$; de (4-20) on déduit : $\lim_{n \to \infty} p[\theta_n \geqslant \text{Log } A] = 1$.

L'évènement "$\theta_n > \text{Log } A$" peut être analysé ainsi :

"$\theta_n > \text{Log } A$" = " $N < n$ " et "$\theta_n - \theta_N > o$"

passant aux probabilités :

$p[\theta_n \geqslant \text{Log } A] = p[N < n].\ p[\theta_n - \theta_N \geqslant o]$

donc on obtient en passant aux limites :

$$\lim_{n \to \infty} p[N < n] \geqslant \lim_{n \to \infty} p[\theta_n \geqslant \text{Log } A] = 1$$

soit $p[N < \infty] = 1$ (4-22)

Le raisonnement est le même pour $E_\zeta\ \theta < o$

IV. 4. Valeur moyenne du nombre d'observations nécessaire pour détecter.

Dans ce paragraphe, nous allons voir quelle est la sensibilité du nombre moyen d'observations nécessaire pour détecter aux variations du modèle de y(t).

IV. 4. 1. Calcul du nombre moyen d'observations.

Supposons que le test se termine à l'étape N, on doit avoir :

$$E_\zeta[\theta_n / n = N] = L(a)\ \text{Log}\ B + (1-L(a))\ \text{Log}\ A$$

où L(a) est la probabilité d'accepter H_o quand a est vraie

N, point d'arrêt du test est une variable aléatoire dépendant de a : N = n(a)

Nous avons démontré dans le § IV. 2. que $E_\zeta\ \theta_n$, (n fixe) est une fonction linéaire de n.

De ce résultat on déduit :

$$E_a[\theta_n / n = N] = E[n(a)] \cdot E_\zeta\ \theta$$

soit :

$$(4\text{-}23) \qquad E[n(a)] = \frac{L(a)\text{Log}\ B + (1-L(a)\ \text{Log}\ A}{E_\zeta\ \theta}$$

E[n(a)] est le nombre moyen d'observations nécessaire pour détecter.

IV. 4. 2. Sensibilité de E[n(a)] aux variations du modèle.

Si $a = a_o$, nous savons que : $E[n(a_o)] = \dfrac{(1-\alpha)\text{Log}\ B + \alpha\ \text{Log}\ A}{E_o\ \theta}$

De même si $a = a_1$: $E[n(a_1)] = \dfrac{\beta\text{Log}\ B + (1-\beta)\ \text{Log}\ A}{E_1\ \theta}$

Si $a \neq a_o$ et $a \neq a_1$, nous savons d'après (4-23) que E[n(a)] s'exprime en fonction de L(a) et $E_\zeta\ \theta$.

La méthode de calcul de L(a) est explicitée à l'appendice [I]. Mais il est difficile d'exprimer $h_o(a_o, a_1, a)$ vérifiant $E\ e^{h_o \theta_n} = 1$, les calculs étant très complexes.

Néanmoins, nous connaissons la forme typique de la représentation de la fonction L(a), en particulier :

pour $a < a_o < a_1$, $L(a) \geqslant 1 - \alpha$

pour $a > a_1 > a_o$, $L(a) \leqslant \beta$.

En tenant compte de cette information, nous allons chercher une valeur approchée du nombre moyen d'observations.

1°) Si $a < a_o < a_1$, ce qui correspond à $\zeta < o$

Nous savons alors que : $1 - \alpha \leqslant L(a) < 1$

Donc $E[n(a)]$ est telle que :

$$(4\text{-}24) \qquad \frac{\text{Log } A + (1-\alpha)(\text{Log } B - \text{Log } A)}{E_\zeta\, \theta} \leqslant E[n(a)] \leqslant \frac{\text{Log } B}{E_\zeta\, \theta}$$

On peut aussi l'écrire :

$$(4\text{-}24\text{bis}) \qquad \frac{\text{Log } B}{E_\zeta\, \theta} - \alpha\, \frac{\text{Log } B - \text{Log } A}{E_\zeta\, \theta} \leqslant E[n(a)] \leqslant \frac{\text{Log } B}{E_\zeta\, \theta}$$

Cet intervalle donne une bonne approximation de $E[n(a)]$:

En effet si $\bar{N} = \dfrac{\text{Log } B}{E_\zeta\, \theta}$, et si nous prenons $\alpha = \beta$, donc $A = 1/B$, alors nous obtenons :

$$(4\text{-}25) \qquad \bar{N}(1-2\alpha) < E(n(a)) < \bar{N}$$

<u>$E[n(a)]$ est connue à 2α °/₀ près</u>.

Nous pourrons alors considérer que $E[n(a)]$ varie comme $\bar{N}$.

Compte tenu du théorème 4-4, <u>plus a s'éloigne de a_o avec $a < a_o$, plus le nombre moyen d'observations décroît</u>.

En effet,

$\exists\, \gamma_o$ tel que pour $\zeta < \gamma_o$ on ait : $\dfrac{\text{Log } B}{E_\zeta\, \theta} < \dfrac{\text{Log } A + (1-\alpha)(\text{Log } B - \text{Log } A)}{E_o\, \theta}$

alors quand ζ décroît avec $\zeta < \gamma_o$, $E[n(a)]$ décroît avec $E[n(a)] < E[n(a_o)]$.

2°) Si $a > a_1 > a_o$, correspond à $\zeta > 1$

On a : $o < L(a) < \beta$ et $E_\zeta\ \theta > o$

Il vient :

$$(4\text{-}26) \qquad \frac{\text{Log } A + \beta(\text{Log } B - \text{Log } A)}{E_\zeta\ \theta} < E[N(a)] < \frac{\text{Log } A}{E_\zeta\ \theta}$$

Comme précédemment, cet intervalle donne une bonne approximation de $E[n(a)]$.
Pour $\alpha = \beta$, c'est-à-dire $A = 1/B$, E[n(a)] est connue à 2 β % près.
Compte tenu du théorème 4-4, plus a s'éloigne de a_1 avec $a > a_1$, plus le nombre moyen d'observations décroît.

En effet :

$\exists\ \gamma_1$ tel que pour $\zeta > \gamma_1$, on ait : $\dfrac{\text{Log } A}{E_\zeta\ \theta} < \dfrac{\text{Log } A + \beta(\text{Log } B - \text{Log } A)}{E_1\ \theta}$

donc, pour ζ croissant avec $\zeta > \gamma_1$, $E[n(a)]$ décroît avec $E[n(a)] < E[n(a_1)]$.

3°) Si $a_o < a < a_1$ correspond à $o < \zeta < 1$

Nous avons vu que pour $a_o < a < a_1$, la probabilité d'erreur peut être importante :
Soit ζ' la valeur de ζ pour laquelle $E_{\zeta'}\theta = o$.
On a bien : $o < \zeta' < 1$ puisque $E_\zeta\ \theta$ est une fonction monotone croissante de ζ
et que : $E_1\ \theta > o$
$E_o\ \theta < o$
Soit $a' = a(\zeta')$
Alors pour $a = a'$, nous savons (appendice [I]) que :

$$(4\text{-}27) \qquad L(a') = \frac{\text{Log } A}{\text{Log } (A/B)}$$

ou encore, pour $\alpha = \beta$ $(A = 1/B)$:

$$(4\text{-}28) \qquad L(a') = \frac{1}{2}$$

on a autant de chances de choisir une hypothèse plutôt que l'autre.

Remarquons que pour $\zeta = \zeta'$, la relation (4-23) est indéterminée, le numérateur et le dénominateur étant nul.

Afin de trouver la valeur de $E[n(a)]$, considérons $E\,\theta_N^2$:

$$E\,\theta_N^2 = L(a')(\text{Log } B)^2 + (1-L(a'))(\text{Log } A)^2$$

Compte tenu de (4-27),

$$E\,\theta_N^2 = -\text{Log A Log B}$$

Or :

$$E\,\theta_N^2 = \text{var } \theta_N \quad (\text{puisque } E_\zeta\, \theta_N = o)$$

$$\text{var } \theta_N \leqslant N\,\nu$$

on a donc l'inégalité :

$$E[n(a')] \geqslant -\frac{\text{Log A Log B}}{\nu} \tag{4-29}$$

où $\nu = \frac{1}{4}\left(\frac{\sqrt{\gamma_1}}{\sigma_1^2} + \frac{\sqrt{\gamma_o}}{\sigma_o^2}\right)^2$

Le résultat de la détection étant peu significatif pour $a_o < a < a_1$, et le nombre d'observations nécessaire pour détecter étant aléatoire, nous allons chercher dans le paragraphe suivant à caractériser, l'ensemble des valeurs de a telles que $a \in]a_o, a_1[$.

IV. 5. Troncature du test de détection.

Nous avons vu que le résultat de la détection quand a se trouve entre a_o et a_1 n'est pas significatif.

Il est alors suffisant et tout aussi utile de savoir non pas si a est proche de a_o ou de a_1 avec une grande probabilité d'erreur, mais si a se trouve entre a_o et a_1 avec une faible probabilité d'erreur.

La règle de troncature que nous présentons, tout en limitant le nombre d'observations, permet de caractériser la zone entre a_o et a_1.

IV. 5. 1. Règle de troncature.

Pour résoudre ce problème, nous adoptons la règle de troncature suivante : On poursuit le test jusqu'à ce qu'une décision soit prise ou que l'étape N soit atteinte.

A l'étape N, si aucune décision n'a été prise le test est arrêté et nous décidons que a se trouve entre a_o et a_1 avec une probabilité d'erreur donnée.
Sachant : $a = \zeta\, a_1 + (1-\zeta)a_o$, il reste à déterminer N tel que :
η étant fixé, $\forall n \geqslant N$ et si aucune décision n'a été prise, alors :

$$p[\zeta \in [o,1]] \geqslant 1 - \eta$$

ou encore, sachant que l'application $\zeta \to E_\zeta\ \theta$ est monotone croissante (théorème 4-4),

$$p\{E_\zeta\ \theta \in [E_o\theta,\ E_1\theta]\} \geqslant 1 - \eta$$

Pour évaluer cette dernière probabilité, il faut trouver une bonne estimée de $E_\zeta\ \theta$ à partir des observations.
Elle nous est donnée par le théorème suivant :

Théorème 4-7
Pour n grand, $\frac{\theta_n}{n}$ est asymptotiquement normale et converge vers $E_\zeta\ \theta$.

Démonstration
On peut écrire : $\frac{\theta_n}{n} = \text{Log}\ \frac{\sigma_o}{\sigma_1} - \frac{1}{2}\left[\frac{\hat{C}_{1n}}{\sigma_1^2} - \frac{\hat{C}_{on}}{\sigma_o^2}\right]$

alors $E\left(\frac{\theta_n}{n}\right) = E_\zeta\ \theta$

$\text{var}\left(\frac{\theta_n}{n}\right) \leqslant \frac{v}{n}$ d'après les résultats précédents.

Parzen [11] a montré que pour n grand, $\hat{C}_{1n}$ et $\hat{C}_{on}$ sont asymptotiquement normales (ils appartiennent à la classe des processus linéaires).
$\frac{\theta_n}{n}$ combinaison linéaire de $\hat{C}_{1n}$ et $\hat{C}_{on}$ est asymptotiquement normale.

■

IV. 5. 2. Détermination de N

Posons : $M = \sup\{-E_o\theta,\ E_1\theta\}$

$\text{Log}\ A = -\text{Log}\ B = m$ soit, $\alpha = \beta$

Si aucune décision n'a été prise, alors, $\frac{-m}{n} < \frac{\theta_n}{n} < \frac{m}{n}$

et $\forall t > o \quad \left|\frac{\theta_n}{n} - E_\zeta\ \theta\right| \leqslant t \to |E_\zeta\ \theta| \leqslant \frac{m}{n} + t$

d'où il vient : $p\{|E_\zeta\theta| \leqslant \frac{m}{n} + t\} \geqslant p\{|\frac{\theta_n}{n} - E_\zeta\theta| \leqslant t\}$

ou encore : $p\{|E_\zeta\theta| \geqslant \frac{m}{n} + t\} \leqslant p\{|\frac{\theta_n}{n} - E_\zeta\theta| \geqslant t\}$ (4-30)

d'après l'inégalité de Bienaymé-Tchebitchev

$$p\{|\frac{\theta_n}{n} - E_\zeta\theta| \geqslant t\} \leqslant \frac{\mathrm{var}\left(\frac{\theta_n}{n}\right)}{t^2} \leqslant \frac{\nu}{n.\,t^2} \qquad (4\text{-}31)$$

(4-30) et (4-31) entrainent :

$$p\{|E_\zeta\theta| \geqslant \frac{m}{n} + t\} \leqslant \frac{\nu}{nt^2} \qquad (4\text{-}32)$$

posons :
$$\left.\begin{aligned} \frac{\nu}{nt^2} &= \eta \\ \frac{m}{n} + t &= M \end{aligned}\right\} \quad (4\text{-}33)$$

(4-32) devient : $p\{|E_\zeta\theta| \geqslant M\} \leqslant \eta$ (4-34)

(4-34) n'est autre que la relation cherchée (probabilité d'erreur inférieure à η)
Le système de deux équations (4-33) permet de déterminer N. En éliminant t,
ce système se ramène à une équation du 2è degré :

$$\eta M^2 N^2 - (2mM + \nu)N + \eta m^2 = o \qquad (4\text{-}35)$$

Cette équation à deux racines positives, l'une correspond à une valeur de t négative, l'autre à une valeur de t positive. Nous avons considéré $t > o$. Par conséquent, nous choisirons pour N, la racine la plus grande de l'équation (4-35).

La relation (4-34) est toujours vérifiée pour $n > N$.
En effet, soit $n > N$, (4-32) est toujours vraie.

or $\frac{\nu}{nt^2} \leqslant \eta$ et $\frac{m}{n} + t \leqslant M$

de plus, $|E_\zeta\theta| \geqslant M \rightarrow |E_\zeta\theta| \geqslant \frac{m}{n} + t$

alors : $p\{|E_\zeta\theta| \geqslant \frac{m}{n} + t\} \geqslant p\{|E_\zeta\theta| \geqslant M\}$ (4-36)

(4-32) et (4-36) impliquent :

$$p\{|E_\zeta \theta| \geq M\} \leq \frac{\nu}{nt^2}$$

soit : $$p\{|E_\zeta \theta| \geq M\} \leq \eta \qquad (4\text{-}37)$$

On déduit de tout ceci, la règle pour le choix de N.
La valeur de N correspond à la racine la plus élevée de l'équation (4-35).

IV. 6. Conclusion

Nous avons étudié dans ce chapitre, la façon dont réagit le test de détection entre deux structures gaussiennes markoviennes quand le vrai modèle de y(t) ne correspond à aucun des modèles envisagés.
Cependant, nous nous sommes limités à une certaine forme de variation du modèle de y(t) en fonction des deux modèles envisagés ((o) et (1)).

La question de savoir si le test se termine en un temps fini est d'une grande importance. Wald [15] indique que la probabilité est 1 pour le test se termine en un temps fini quand les observations sont indépendantes.
Nous avons montré dans ce chapitre que quel que soit le vrai modèle, le test entre deux structures gaussiennes markoviennes se termine presque surement en un temps fini.

En faisant une hypothèse à propos du régime transitoire du filtre de Kalman, nous avons été en mesure de donner une expression simple du nombre moyen d'observations nécessaire au test pour prendre une décision dans le cas général.

Le calcul de la probabilité d'erreur quand le vrai modèle n'est ni (o) ni (1), est difficile à expliciter. Néanmoins, nous savons que lorsque le vrai modèle se trouve "entre" les modèles (o) et (1), le résultat de la détection n'est pas significatif alors que cette probabilité d'erreur est faible lorsque le vrai modèle est à "l'extérieur" des deux modèles (o) et (1).
Nous avons alors cherché à caractériser la zone se trouvant "entre" les modèles (o) et (1) par une règle de troncature ; le nombre d'observations étant aléatoire et pouvant devenir très grand.

Chapitre V

Détection de changement des caractéristiques d'un processus à représentation gaussienne-markovienne.

Nous étudions dans ce chapitre le problème de la détection de changement de paramètre :
La valeur initiale est connue : $a = a_o$
Mais à la différence des deux premiers chapitres, la valeur finale après le changement n'est pas connue ; on sait seulement qu'elle appartient à un certain ensemble de valeurs.

Soit Ω_1 cet ensemble. Il est caractérisé de la façon suivante :
Soient a' et a" deux éléments de Ω_1 ; a' et a" sont alors liés par la relation suivante :

(5-1) $\quad a" = \zeta\, a' + (1-\zeta)\, a_o \quad , \quad \zeta \in R.$

De plus, on suppose que les éléments de Ω_1 sont tels que toutes les conditions nécessaires de stabilité ou de positivité soient satisfaites.

Les algorithmes de détection présentés dans les deux premiers chapitres étaient basés sur un test d'hypothèses "simples" : $a = a_o$ ou $a = a_1$.
Ici, les hypothèses considérées sont des hypothèses "multiples" : $a = a_o$ ou $a \in \Omega_1$.
Nous allons analyser l'ensemble des tests séquentiels possibles pour ces deux hypothèses, puis nous étudierons les deux types de changement suivants :
- détection de changement brusque de paramètre au moment où il arrive, il s'agit de détecter le changement le plus rapidement possible, mais à la différence du chapitre III, seul, l'intervalle auquel appartient la valeur finale est connue.

- détection de variation de paramètre ; on teste si le paramètre a varié de façon sensible ou non. (La valeur de la variation n'est pas connue, seule sa "direction" est connue).

V. 1. Choix d'un test séquentiel d'hypothèses "multiples".

1°) Construction des tests possibles.

Les deux hypothèses à tester sont :

$$\mathcal{H}_0 \qquad a = a_0$$
$$\mathcal{H}_1 \qquad a \in \Omega_1$$

Nous sommes en présence d'hypothèses "multiples". (hypothèse $\mathcal{H}_1$)

Nous utilisons les mêmes notations qu'au chapitre II ;

Soit :

Ω_0 l'ensemble constitué par $a = a_0$

on suppose : $\Omega_0 \cap \Omega_1 = \emptyset$ c'est-à-dire : $a \notin \Omega_1$

D'après le § II. 2. 5., s'il est possible de définir une fonction non négative $p_1(a)$ telle que :

$$\int_{\Omega_1} p_1(a)\, da = 1$$

alors, le test séquentiel de $\mathcal{H}_0$ et $\mathcal{H}_1$ sera basé sur le rapport de vraisemblance :

$$(5\text{-}2) \qquad \Lambda(\mathcal{Y}_n) = \frac{\int_{\Omega_1} p_1(a)\, p[\mathcal{Y}_n/a]\, da}{p[\mathcal{Y}_n/a_0]}$$

et l'algorithme de test sera le même qu'en II. 2. 5.

Il est possible de construire autant de tests séquentiels que de fonctions $p_1(a)$ existantes.

2°) Critère d'optimalité

Soient : α = probabilité de choisir $\mathcal{H}_1$ sachant $\mathcal{H}_0$

$\beta(a)$ = " " $\mathcal{H}_0$ " a

$$\bar{\beta} = \int_{\Omega_1} p_1(a)\beta(a)\, da$$

A et B étant choisies, α et $\bar{\beta}$ sont déterminées par les relations :

$$\alpha = \frac{1 - B}{A - B} \quad , \quad \bar{\beta} = \frac{B(A-1)}{A - B}$$

Ainsi, le choix de $p_1(a)$ affecte et affecte seulement la probabilité d'erreur $\beta(a)$.
Une fonction $p_1(a)$ peut être considérée comme d'autant plus intéressante que la valeur maximale de $\beta(a)$ correspondante est plus petite.
On peut alors énoncer un critère d'optimalité :

Définition 5. 1.

Parmi l'ensemble des fonctions $p_1(a)$ existantes, la fonction optimale sera celle qui minimise $\max_{a \in \Omega_1} [\beta(a)]$

Il n'y a aucune méthode générale pour la détermination de la fonction optimale. Mais nous pouvons énoncer le théorème suivant du à Wald [15].

Théorème 5. 1.

Soit S_1 la frontière de Ω_1

S'il existe une fonction $v_1(a)$ non négative définie sur S_1 telle que :

$$(5\text{-}3) \qquad \int_{S_1} v_1(a)\, ds = 1$$

$$(5\text{-}4) \quad \text{Le test basé sur le rapport de vraisemblance } \Lambda(\mathcal{Y}_n) = \frac{\int_{S_1} v_1(a) p[\mathcal{Y}_n/a]\, ds}{p[\mathcal{Y}_n/a_o]}$$

satisfait aux deux conditions : $\begin{cases} \beta(a) = \beta & a \in S_1 \\ \beta(a) \leqslant \beta & a \in \Omega_1 \end{cases}$

alors, cette fonction $v_1(a)$ est une fonction optimale dans le sens défini au-dessus.

Démonstration.

$v_1(a)$ définie sur S_1 peut être considérée comme cas limite d'une fonction $p_1(a)$ nulle pour tout point a de Ω_1 dont la distance est supérieure à Δ, $\Delta \to o$.
Des conditions précédentes, il s'ensuit : $\max_{a \in \Omega_1} \beta(a) = \beta = \int_{S_1} v_1(a)\, \beta(a) ds = \bar{\beta}$

Montrons que $v_1(a)$ est optimale :

Soit $p_1^*(a)$, une fonction non négative définie sur Ω, et $\beta^*(a)$ la probabilité d'erreur correspondante.
D'après les résultats précédents :

$$\int_{\Omega_1} p_1^*(a)\, \beta(a)\, da = \int_{\Omega_1} v_1(a)\, \beta(a)\, da = \bar{\beta}$$

Il s'ensuit que : $\max_{a \in \Omega_1} \beta^*(a) \geqslant \bar{\beta}$

$v_1(a)$ est donc meilleure que $p_1^*(a)$, puisque $\max_{a \in \Omega_1} \beta(a) = \bar{\beta}$

■

3°) Test séquentiel optimum d'hypothèses "multiples".

Ainsi, si l'on peut trouver une fonction $v_1(a)$ définie sur la frontière S_1 de Ω_1 telle que :

$$(5\text{-}5) \quad \begin{cases} \displaystyle\int_{S_1} v_1(a)\, ds = 1 \\ \\ \text{le test basé sur } v_1(a) \text{ est tel que } \begin{cases} \beta(a) = \beta \\ a \in S_1 \end{cases} \\ \qquad\qquad\qquad\qquad\quad \begin{cases} \beta(a) \leqslant \beta \\ a \in \Omega_1 \end{cases} \end{cases}$$

alors le test correspondant sera optimum au sens précisé dans la définition 5.1.

Cette analyse sur le choix d'un test séquentiel pour les deux hypothèses $\mathcal{H}_0$ $(a = a_0)$ et $\mathcal{H}_1$ $(a \in \Omega_1)$ va nous permettre de traiter les deux problèmes de détection abordés dans les § suivants.

V. 2. Détection de changement brusque de paramètre.

Dans le chapitre III , le paramètre a pouvait varier brusquement de a_0 à a_1.

Mais généralement, en pratique, la valeur finale n'est pas connue.
Dans le meilleur des cas, seul, l'intervalle de changement est connu.

Soit Γ l'intervalle de changement :

$\Gamma = [a_1 - \Delta a_1, a_1 + \Delta a_1]$ (alors $\Gamma \subset \Omega(a_1, a_o)$) (ensemble défini au chapitre IV)

où Δa_1 est tel que : $\Delta a_1 = \mu(a_1 - a_o)$. $\mu \in R$

Dans le problème que nous traitons, le paramètre a est susceptible de passer brusquement de a_o à une valeur appartenant à Γ.
Nous cherchons à détecter ce changement aussi rapidement que possible.

Pour cela, nous allons employer la même procédure de détection que celle exposée dans le second chapitre :
On fait plusieurs tests de façon continue jusqu'à détection du changement, chaque test utilisant le résultat du test précédent.
Nous allons construire un test permettant de détecter si $a = a_o$ ou si $a \in \Gamma$.

V. 2. 1. Construction du test de détection d'hypothèses multiples.

Soient les deux hypothèses :

$\mathcal{H}_o$: $a = a_o$

$\mathcal{H}_1$: $a \in \Gamma$

Pour traiter ces deux hypothèses, nous allons utiliser un test séquentiel d'hypothèses "multiples" :
Précisons les notations :
Soient : Ω_o l'ensemble constitué par $a = a_o$

Ω_1 " " " $a \geqslant a_1 - \Delta a_1$; alors $\Gamma \subset \Omega_1$

Nous avons vu dans le § V. 1. que l'on pouvait définir autant de tests que de fonctions $p_1(a)$ existantes définies sur Ω_1.
Pour notre cas particulier nous ne considérerons que le sous-ensemble des fonctions $p_1(a)$ telles que $p_1(a) = o$ pour $a \geqslant a_1 + \Delta a_1$; les seules valeurs possibles de a étant a_o et $a \in [a_1 - \Delta a_1, a_1 + \Delta a_1]$.

Parmi ce sous-ensemble, s'il existe une fonction vérifiant les conditions (5-5), nous avons vu que le test correspondant était optimum.

Une telle fonction existe :

Théorème 5.2.

La fonction $v_1(a) = \delta(a-a_1+\Delta a_1)$ où δ est l'impulsion de Dirac, vérifie l'ensemble des conditions (5-5).

Démonstration

La surface S_1 de Ω_1 est constituée du point $a_1 - \Delta a_1$.
$v_1(a)$ est définie sur S_1 et non négative.

De plus on a : $\int_{S_1} v_1(a)\, ds = 1$

Le test basé sur le rapport : $\Lambda(\mathcal{Y}_n) = \dfrac{p[\mathcal{Y}_n/a_1-\Delta a_1]}{p[\mathcal{Y}_n/a_o]}$ converge

pour toutes les valeurs de $a \in \Omega_1$ (théorème 4-6).
Nous avons : $\beta(a) = L(a)$.
Or pour $a > a_1 - \Delta a_1 > a_o$, le test basé sur $\Lambda(\mathcal{Y}_n)$ défini au-dessus est tel que : $o \leqslant L(a) \leqslant \beta$ (d'après le chapitre IV).
Donc pour $a > a_1 - \Delta a_1$, on a : $\beta(a) \leqslant \beta(a_1) = \beta$.
Les conditions (5-5) sont remplies.
La fonction $v_1(a)$ ainsi définie vérifie en outre : $v_1(a) = o$ pour $a > a_1 + \Delta a_1$.
Le test basé sur cette fonction est donc optimum au sens de la définition 5.1. ■

V. 2. 2. Algorithme de détection de changement par saut.

L'algorithme de détection de changement par saut sera équivalent à celui proposé dans le second chapitre.
Le test séquentiel d'hypothèses "multiples" utilisé est un test simple de détection entre deux structures gaussiennes markoviennes ($a = a_o$ et $a = a_1 - \Delta a_1$)
Ce test est basé sur le rapport de vraisemblance :

$$\Lambda(\mathcal{Y}_n) = \frac{p[\mathcal{Y}_n/a_1-\Delta a_1]}{p[\mathcal{Y}_n/a_o]}$$

L'algorithme demandera un temps et un volume de calcul pratiquement similaire à celui du chapitre III.

Le calcul des performances se fait pratiquement de la même façon ;

Soient :

T_b : le temps moyen nécessaire pour atteindre l'une des bornes au début de l'algorithme.

T_a : le temps de retard moyen de détection de changement.

Nous avons les relations :

$$(5\text{-}6) \qquad \begin{aligned} T_b &= E[n(a_o)] \cdot \Delta T \\ T_a &= E[n(a)] \cdot \Delta T \end{aligned}$$

où $a \in [a_1 - \Delta a_1, a_1 + \Delta a_1]$ ΔT = période d'échantillonnage.

avec :

$$(5\text{-}7) \qquad E[n(a_o)] = \frac{\alpha \operatorname{Log} 1 + (1-\alpha) \operatorname{Log} B}{E_o \theta}$$

$$(5\text{-}8) \qquad E[n(a)] \leqslant \frac{\operatorname{Log} A - \operatorname{Log} B}{E_\zeta \theta} \qquad \zeta \text{ correspond à } a.$$

T_a et T_b, s'expriment en fonction des probabilités d'erreur α, β. Nous savons de plus que la probabilité d'erreur de détection de changement $\beta(a)$ $(a \in \Gamma)$ est telle que $\beta(a) \leqslant \beta$.

On peut donc, suivant ce que l'on cherche, rapidité de détection ou exactitude de la détection, se fixer les probabilités d'erreur maximales, donc les bornes A et B.

V. 3. Détection de variation de paramètre.

Le problème que nous traitons maintenant est différent du problème précédent :

Dans le paragraphe précédent, le paramètre était susceptible de changer par saut.

On cherchait alors à détecter l'instant τ du changement aussi rapidement que possible.

Dans ce paragraphe, le paramètre est susceptible de varier de façon continue ou discontinue. A un certain moment, on teste si le paramètre a varié de façon sensible ou non.

Soit a_o la valeur initiale du paramètre.

Le paramètre est susceptible de varier suivant une direction donnée :

Etant donné un point a' (caractérisant la direction), la valeur finale a est telle que $a \geqslant a_o$ avec a lié à a' et a_o par

$$a = \zeta\, a' + (1-\zeta)\, a_o$$

nous cherchons alors à tester si : $a = a_o$ ou $a > a_o$

(on pourrait également considérer le cas $a < a_o$).

V. 3. 1. Position du problème.

Si a est très proche de a_o, il est raisonnable, du moins en pratique de considérer que le paramètre n'a pas changé. On peut donc affecter un degré de préférence croissant pour la détection de variation, au fur et à mesure que a s'éloigne de a_o.

Il est alors possible de trouver une valeur a_1, $(a_1 = \zeta_1\, a' + (1-\zeta_1)a_o)$ telle que :

si $a = a_o$, le modèle n'a pas changé

si $a \geqslant a_1$, le modèle est considéré comme défectueux, il est important de détecter la variation.

si $a_o < a < a_1$, a est proche de a_o, la détection de variation n'a pas d'importance pratique.

Ayant choisi a_1, les hypothèses à tester sont :

$\mathcal{H}_o$: $a = a_o$

$\mathcal{H}_1$: $a \geqslant a_1$

L'intervalle $]a_o, a_1[$ est une zone d'indifférence pour la détection de changement.

L'algorithme de détection de variation revient à tester les hypothèses "multiples" $\mathcal{H}_o$ et $\mathcal{H}_1$.

V. 3. 2. Construction du test de détection d'hypothèses multiples $\mathcal{H}_o, \mathcal{H}_1$.

Soient : Ω_1 l'ensemble des valeurs de a tel que $a \geqslant a_1$

S_1 la frontière de Ω_1

S_1 est constitué du point $a = a_1$

Il y a autant de tests que de fonctions $p_1(a)$ définies sur Ω_1.

Mais il existe une fonction optimale au sens de la définition 5-1 :

Théorème 5-3

La fonction $v_1(a) = \delta(a-a_1)$, δ impulsion de Dirac, vérifie l'ensemble des conditions (5-5).

Démonstration.

Elle est équivalente à celle du théorème 5-2 et s'appuie sur les résultats du chapitre IV.

Le test de détection optimum est donc basé sur le rapport de vraisemblance :

$$\Lambda(\mathcal{Y}_n) = \frac{p[\mathcal{Y}_n/a_1]}{p[\mathcal{Y}_n/a_o]}$$

V. 3. 3. Algorithme de détection de variation.

L'algorithme de détection de variation du paramètre a est alors basé sur un test simple de détection entre deux structures gaussiennes markoviennes ($a = a_o$ et $a = a_1$).

Mais, ici, le paramètre a peut prendre des valeurs différentes de a_o et a_1.

Le volume et le temps de calcul nécessaire pour l'algorithme est donc pratiquement équivalent à celui présenté au chapitre II.

1°) Comportement de l'algorithme.

L'analyse faite au chapitre IV sur la sensibilité du test simple aux variations du paramètre a ($a \neq a_o$ et $a \neq a_1$), va nous permettre d'en déduire le comportement de l'algorithme de détection de variation.

Quelque soit la valeur de a_1 l'algorithme prend une décision en un temps fini.

quand $a = a_o$:

La probabilité d'erreur de détection est α.

Le temps moyen d'observation T_{ob} est :

(5-9) $$T_{ob} = \Delta T.\ E[n(a_o)]$$

où $$E[n(a_o)] = \frac{\alpha \text{Log } A + (1-\alpha)\text{Log } B}{E_o\ \theta}$$

quand $a \geqslant a_1$

La probabilité d'erreur de détection est telle que $\beta(a) \leqslant \beta$.

Le temps moyen d'observation T_{ob} est

$$T_{ob} = \Delta T.\ E[n(a)]$$

où $E[n(a)] < \frac{\text{Log } A}{E_\zeta\ \theta}$, ζ correspond à a.

c'est-à-dire :

(5-10) $$T_{ob} < \frac{\Delta T.\ \text{Log } A}{E_{\zeta_1}\ \theta}$$, ζ_1 correspond à a_1

2°) Caractérisation de la zone d'indifférence.

Quand $a_o < a < a_1$, nous avons vu dans le précédent chapitre que le résultat est peu significatif.

La probabilité d'erreur peut varier de β à $1-\alpha$.

Le temps d'observation peut également devenir très grand.

Nous avons alors proposer une règle de troncature. Cette règle de troncature permet de caractériser la zone d'indifférence :

Soit $T_t = N.\ \Delta T$ où N est déterminé par l'équation (4.3 5). Si au bout d'un temps $\tau = T_t$ aucune décision n'a été prise, alors l'algorithme est arrêté, et nous sommes certain, avec une probabilité d'erreur η fixée, que le paramètre a appartient à la zone d'indifférence.

L'algorithme se termine au plus au bout d'un temps T_t.

3°) Choix de a_1

L'algorithme de détection de variation repose sur le choix de a_1. Ce choix peut se faire suivant des considérations physiques sur le processus technologique observé par y(t). Il peut se faire également suivant des considérations

de rapidité de détection :
Plus a_1 est proche de a_0 et plus, le test sera sensible aux faibles variations de a, mais la détection des grandes variations de a sera moins rapide. Ce problème est abordé dans le chapitre suivant.

V.4. Conclusion.

Nous venons d'aborder deux types de problèmes de changement de paramètres dans lesquels la valeur finale n'est pas connue.
Les hypothèses à tester sont des hypothèses "multiples" ($\mathcal{H}_1$), plusieurs valeurs étant à considérer. Il existe de nombreuses solutions possibles pour lesquelles en général l'expression du rapport de vraisemblance est complexe.

L'étude sur la sensibilité du détecteur entre deux structures gaussiennes markoviennes est un théorème du à Wald nous ont permis de trouver des solutions optimales (au sens précisé dans la définition 5-1) pour ces deux problèmes. Outre l'optimalité, il faut souligner la simplicité des algorithmes présentés : L'expression du rapport de vraisemblance ne met en jeu que deux valeurs du paramètre (bien que plusieurs valeurs soient possibles).
Ces algorithmes demandent alors un temps et un volume de calcul pratiquement équivalent au test entre deux structures gaussiennes markoviennes.
Pour le second algorithme, il est nécessaire de choisir une valeur a_1 telle que si $a \geqslant a_1$ le modèle soit considéré comme défectueux.

Ces deux algorithmes ont été programmés, les expérimentations sont présentées dans le chapitre suivant.

Chapitre VI

Etude expérimentale des algorithmes proposés

Dans le chapitre précédent, deux algorithmes sont proposés pour la détection de changement brusque ou de variation de paramètre.

Ces deux algorithmes vont être simulés en temps réel à l'aide d'une installation de calcul hybride (calculateur analogique + calculateur numérique). Les signaux $y(t)$ sont générés à partir du calculateur analogique. Après discrétisation, les observations sont traitées séquentiellement par le calculateur numérique. L'évolution du rapport de vraisemblance, θ_n, est visualisée sur un écran cathodique.

Nous présentons également une étude des performances des algorithmes :
Le critère de performance utilisé est :
- le temps de retard moyen pour l'algorithme de détection de changement par saut (temps entre l'instant τ du changement et l'instant τ' de détection du changement).
- le temps moyen d'observation nécessaire pour détecter à l'algorithme de détection de variation.

Pour le second algorithme, nous aborderons le problème du choix de a_1.

VI. 1. Simulation en temps réel des algorithmes

VI. 1. 1. Simulation et modèle du signal à représentation gaussienne markovienne y(t).

1°) Mise en oeuvre du signal y(t).

Le signal y(t) est généré sur un calculateur analogique (EAI 380). Pour ce faire, un système linéaire est cablé sur le calculateur. A l'entrée de ce système est branché un générateur de"bruit blanc gaussien". (Hewlett Packard type 3722). Ce générateur délivre un bruit pseudoblanc avec des séquences de longueur finie ou infinie et une fréquence de coupure choisie par l'utilisateur.

La fréquence de coupure choisie pour toutes les simulations est de 5 KHZ. D'autre part, les séquences infinies sont préférées aux séquences finies. Le bruit ainsi généré est blanc avec une bonne approximation. La sortie du système linéaire est bruitée :

$$y(t) = y'(t) + w(t)$$

Un autre générateur de "bruit blanc gaussien" est donc branché à la sortie du système linéaire.

Nous avons choisi deux systèmes pour les expérimentations : un système du premier ordre et un système du second ordre. Les schémas analogiques des systèmes choisis sont présentés en annexe [1]. Nous supposons que les signaux aléatoires ainsi générés sont à représentation gaussienne markovienne avec une bonne approximation (voir GAUTHIER[14]). Les signaux gaussiens markoviens sont transmis au calculateur numérique (T2000) par un bornier. L'échantillonnage est commandé par une horloge interne au calculateur : le cadenceur. Chaque top d'horloge déclenche une conversion analogique digitale dont la fin déclenche à son tour une interruption prioritaire dans le calculateur numérique. Les signaux sont échantillonnées à 10 msec.

2°) Obtention du modèle discret de y(t).

Nous connaissons les équations continues du système linéaire cablé sur le calculateur analogique. Mais le traitement des équations continues sur le calculateur numérique à notre disposition peut devenir très important en temps et en volume de calcul.
Compte tenu de cette contrainte, nous allons utiliser pour le traitement numérique des modèles discrets, ce qui permet de choisir des fréquences d'échantillonnage plus élevées.

Pour obtenir ce modèle, nous utilisons un programme d'identification stochastique qui nous donne directement les équations du filtre correspondant [17]. Dans un premier temps, ce programme estime la fonction d'autocorrélation de y(t). Puis, après factorisation de la matrice de Hankel due à Rissanen [3] et résolution d'une

équation de Riccati discrète (FAURRE [5]), il imprime les matrices : Φ, T, H, R_* vérifiant :

$$(6\text{-}1) \qquad \begin{aligned} X_{k+1} &= \Phi X_k + T W_k \\ y_k &= HX_k + W_k \qquad \text{avec } EW_k^2 = R_* \end{aligned}$$

3°) Paramètre susceptible de varier ou de changer brusquement.

Le paramètre a susceptible de varier ou de changer brusquement est la matrice de dynamique F du système linéaire continu associé à y(t).

Nous avons bien sur la relation :

$$\Phi = e^{F\Delta t}$$ Δt étant la période d'échantillonnage.

Nous avons choisi F car les constantes de temps du système se règlent à partir des potentiomètres sur le calculateur analogique.

On considère les deux cas suivants :

- changement brusque :

Sur le calculateur analogique sont cablés deux systèmes linéaires de même nature, excités par le même bruit blanc gaussien, et dont la sortie est bruitée de la même façon. Seule diffère la valeur d'un potentiomètre intervenant dans la matrice de dynamique F.

On obtient deux signaux :

$y_o(t)$ de matrice de dynamique F_o

$y(t)$ de matrice de dynamique F telle que $F \in [F_1 \pm \Delta F_1]$, F_1 fixée.

Les deux signaux sont branchés au bornier du calculateur numérique.

Celui-ci détermine lui-même l'instant τ du changement :

si $t < \tau$ $\quad y(t) = y_o(t)$

si $t > \tau$ $\quad y(t) = y(t)$

- variation :

On cable un système correspondant à une matrice de dynamique F_o. On fait varier F en changeant la valeur d'un potentiomètre.

VI. 1. 2. Mise en oeuvre numérique des algorithmes.

Les deux algorithmes de détection de changement ou de variation sont basés sur le calcul du rapport de vraisemblance :

$$\theta_n = \text{Log} \frac{p[\mathcal{Y}_n/F_1]}{p[\mathcal{Y}_n/F_o]}$$

Dans le chapitre II et les chapitres suivants, nous avons indiqué la façon de calculer θ_n quand le modèle de y(t) est continu et les observations discrètes. Nous avons déjà indiqué que pour la simulation en temps réel, des modèles discrets seront utilisés, ceci afin d'augmenter la rapidité du temps de calcul. Voyons comment calculer θ_n à partir de modèles discrets.

1°) Calcul de θ_n à partir de modèles discrets.

Nous avons :

$$\theta_n = \theta_{n-1} + \text{Log} \frac{\sigma_o}{\sigma_1} - \frac{1}{2}\left[\frac{\varepsilon_{1n}^2}{\sigma_1^2} - \frac{\varepsilon_{on}^2}{\sigma_o^2}\right]$$

où :

$$\varepsilon_{in} = y(t_n) - \hat{y}_i(t_n) \quad (i = o,1), \quad \hat{y}_i(t_n) = E\{y(t_n)/y(t_o)..y(t_{n-1}),H_i\}$$

Le programme d'identification stochastique nous a donné directement les équations des filtres discrets correspondants aux modèles (o) et (1). (hypothèses H_o et H_1) :

$$X_{k+1} = \Phi_i X_k + T_i W_{k,i}$$

$$y_k = H X_k + W_{k,i} \qquad E W_{k,i}^2 = R_{*,i} \quad (i = o,1)$$

Nous avons alors les relations :

$$\varepsilon_{i,n} = W_{n,i}$$

$$\sigma_i^2 = R_{*,i}$$

$\varepsilon_{i,n}$ est donc donné par les équations :

$$X_i(n+1/n) = \Phi_i X_i(n/n-1) + T_i[y(t_n) - H X_i(n/n-1)]$$

et :

$$\varepsilon_{i,n} = y(t_n) - H X_i(n/n-1) \qquad (i = o,1)$$

Ces deux expressions sont nécessaires pour l'obtention du rapport de vraisemblance θ_n.

Nous présentons maintenant les programmes correspondant aux deux algorithmes :

2°) Programme de détection de changement brusque.

L'algorithme correspondant reprend les mêmes étapes que l'algorithme présenté au chapitre III (§ III. 2. 2.) en remplaçant F_1 par $F_1-\Delta F_1$.
Le programme de détection correspondant à l'algorithme est déclaré sous le nom SEC DECHT
Il comprend :

α - Entrée des matrices correspondants aux deux systèmes envisagés.

β - initialisation (INIT).

γ - boucle d'attente : le programme est mis en attente jusqu'au moment d'une nouvelle interruption (entrée d'une observation). Dans cette partie du programme, on décide également du moment du changement brusque.

δ - calcul de θ_n (CALCUL)

ε - test : si $\theta_n \geqslant$ Log A, l'algorithme est arrêté avec détection de changement
si $\theta_n \leqslant$ Log B, on retourne dans la boucle d'attente avec $\theta_\varphi = o$.

L'organigramme et le listing de ce programme sont présentés en annexe [2].

3°) Programme de détection de variation de paramètre.

L'algorithme est plus simple que celui de changement brusque. Il est similaire à celui décrit au chapitre II (§ II. 4. 3.).
F_1 ayant été choisi, le programme comprend les mêmes parties que le programme précédent. Mais ici, dans la boucle d'attente, on n'a pas à changer de signal.
La partie test est différente :
si $\theta_n \geqslant$ Log A, le programme imprime détection de variation
si $\theta_n \leqslant$ Log B, " " " de non variation.
L'organigramme et le listing sont présentés en annexe [2]. Ce programme est déclaré sous le nom : SEC DETECT.

VI. 1. 3. Résultats.

L'annexe [3] présente des courbes de visualisation du rapport de vraisemblance θ_n.

Pour l'algorithme de détection de changement brusque, le changement de signal est visualisé par un trait vertical. Nous remarquons également une dispersion élevée pour le nombre d'observations nécessaires pour détecter.

VI. 2. Etude des performances des algorithmes proposés.

Nous présentons dans ce paragraphe une étude expérimentale sur le temps de retard moyen (détection de changement brusque) et sur le temps moyen d'observation (détection de variation). Nous allons examiner quelle est la variation de ces critères en fonction des différents facteurs propres au système (covariance du bruit de mesure R et période d'échantillonnage des observations). Regardons d'abord l'expression du temps d'observation correspondant aux différents algorithmes.

VI. 2. 1. Expression du temps d'observation correspondant aux différents algorithmes.

1°) Algorithme de détection de changement par saut (le saut étant connu)

Deux temps d'observations ont été définis :

T_a = temps de retard moyen de détection de changement (temps moyen entre l'instant du changement et l'instant de détection du changement).

T_b = temps moyen nécessaire à l'algorithme pour atteindre l'une des bornes (à partir de l'instant o), sachant qu'il n'y a pas changement.

Les expressions de T_a et T_b sont :

$$T_a = \frac{\text{Log } A - \text{Log } B}{E_1\ \theta} \cdot \Delta T$$

(6-2)

$$T_b = \frac{\alpha \text{Log } A + (1-\alpha)\text{Log } B}{E_o\ \theta} \cdot \Delta T$$

ΔT étant la période d'échantillonnage.

2°) Algorithme de détection de changement par saut (le saut étant inconnu)

De la même façon que dans l'algorithme précédent, nous avons définit :

$$T'_a \text{ et } T'_b$$

Leurs expressions sont :

(6-3)

$$T'_a = \frac{\text{Log } A - \text{Log } B}{E_\zeta\ \theta} \cdot \Delta T \qquad (\zeta \text{ correspond à a, avec } a \in [a_1 - \Delta a_1, a_1 + \Delta a_1]$$

$$T'_b = \frac{\alpha \text{ Log } A + (1-\alpha)\text{Log } B}{E_o\ \theta} \cdot \Delta T$$

3°) Algorithme de détection de variation (variation inconnue)

Supposons a_1 fixé, avec $a_1 > a_o$.

Nous ne considérons que le cas où $a \geqslant a_1 > a_o$. En effet, nous avons vu (IV. 4. 2.) que pour $a_o < a < a_1$, le résultat était peu significatif. Il est donc inutile de considérer ce cas.

Dans le cas où $a \geqslant a_1 > a_o$, l'expression du temps moyen d'observation T_1 est telle que (4-25) :

(6-4)

$$\frac{\text{Log } A + \beta(\text{Log } B - \text{Log } A)}{E_\zeta\ \theta} \cdot \Delta T < T_1 < \frac{\text{Log } A}{E_\zeta\ \theta} \cdot \Delta T$$

4°) Critère de performance étudié.

Il nous a paru utile d'étudier surtout les temps d'observations nécessaires pour décider lorsque un changement ou une variation du paramètre s'était produit.

Dans les expérimentations que nous allons faire, nous prendrons :

$$\alpha = \beta \quad \text{d'où} \quad \text{Log } A = -\text{Log } B$$

Considérons l'expression : $$T = \frac{\text{Log } A}{E_\zeta\ \theta} \cdot \Delta T$$

T_a et T'_a se déduisent facilement de T : $$T_a = 2T(a = a_1)$$

$$T'_a = 2T$$

la valeur de T_1 est connue à $2\beta\,\%$ près : $T(1-2\beta) < T_1 < T$.

T est donc une expression représentative de T_a, T'_a, T_1.
Nous n'étudierons donc qu'un seul critère de performance : le temps d'observation T.

$$(6\text{-}5) \qquad T = \frac{\text{Log } A}{E_\zeta\ \theta} \cdot \Delta T$$

VI. 2. 2. Paramètre susceptible de varier ou de changer et facteurs entrant en jeu dans l'expression de T.

Nous voulons étudier comment se comporte le temps d'observation T pour détecter une variation du paramètre en fonction des différents facteurs entrant en jeu dans l'expression de T.

1°) Paramètre susceptible de varier ou de changer.

Nous avons choisi comme paramètre susceptible de changer par saut ou de varier, la matrice de dynamique F du système linéaire. Elle caractérise en effet les constantes de temps du système physique qui sont les plus susceptibles de varier.

2°) Facteurs entrant en jeu dans l'expression de T.

Les premiers facteurs qui apparaissent dans l'expression de T, sont les probabilités d'erreur α et β. Elles interviennent de façon simple :

$$\text{Log } A = \text{Log } \frac{1 - \beta}{\alpha}$$

Aussi, nous n'avons pas pensé utile de considérer la variation de ces facteurs dans l'étude de T.
Nous avons pris arbitrairement :

$$\alpha = \beta = 0.05.$$

Les autres facteurs intéressants à considérer sont la matrice de covariance du bruit de mesure : R et la période d'échantillonnage : ΔT du signal $y(t)$. Nous examinerons comment varie le temps d'observation T en fonction de ces deux facteurs.

Un dernier facteur intéressant, intervenant dans l'expression de T est le paramètre a_1, (ou F_1). Ce paramètre est déterminé dans le cas d'un changement par saut, mais il reste à fixer pour l'algorithme de détection de variation. Nous examinerons la variation de T en fonction de F_1.

V. 2. 3. Plan d'expérience.

1°) Modèles à représentation gaussienne markovienne utilisés

Deux systèmes vont être utilisés :

- un système du premier ordre
- un système du second ordre.

- Le système du premier ordre a pour équations :

$$\dot{X} = -\lambda X + v \qquad \text{avec } Q = Ev^2 = 0{,}77$$
$$y = x + w \qquad R = Ew^2$$

La covariance du bruit de mesure R varie de 0.0035 à 0.028. La période d'échantillonnage ΔT varie de 2 msec à 16 msec. Les probabilités d'erreur ont pour valeur : $\alpha = \beta = 0.05$.

- Le système du second ordre a pour équations :

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} \lambda & 2 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0.01 \end{bmatrix} \cdot v \qquad \text{avec } Q = Ev^2 = 0.6$$
$$R = Ew^2$$

$$y = [1 \quad 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + w$$

La covariance du bruit de mesure R varie 0.0035 à 0.028. La période d'échantillonnage ΔT varie de 2 msec à 16 msec. Les probabilités d'erreur sont : $\alpha = \beta = 0.05$.

2°) Programmes de calcul de T

L'expression de T fait intervenir le calcul de $E_\zeta\ \theta$. Nous avons vu (§ IV. 1. et IV. 2) comment obtenir $E_\zeta\ \theta$ à partir d'un modèle discret. Ici, nous considérons un modèle continu à observations discrètes.

- Calcul de $E_\zeta\ \theta$.

L'expression de $E_\zeta\ \theta$ est la suivante (16 bis chap. IV)

$$E_\zeta\ \theta = \text{Log}\ \frac{\sigma_o}{\sigma_1} - \frac{1}{2}\left[\frac{C_{o1}}{\sigma_1^2} - \frac{C_{oo}}{\sigma_o^2}\right]$$

où $C_{o1} = E\,\varepsilon_1^2$

$C_{oo} = E\,\varepsilon_o^2$

de plus : $\varepsilon_{in} = y(t_n) - H\,\hat{X}_i(t_{n-})$ $(i = o,1)$, i représente l'indice du modèle utilisé pour le filtre.

Il vient :

$$E\,\varepsilon_{in}^2 = H\,E\{(X(t_n)-\hat{X}_i(t_{n-}))(X(t_n)-\hat{X}_i(t_{n-}))'\}\,H' + R$$

Posons :

$$\Sigma_i = E\{(X(t_n)-\hat{X}_i(t_{n-}))(X(t_n)-\hat{X}_i(t_{n-}))'\}$$

Le calcul de Σ_i en fonction du vrai modèle de y(t) et du modèle i est exposé en appendice [III].

On a donc :

$$(6\text{-}6) \qquad E_\zeta\ \theta = \text{Log}\ \frac{\sigma_o}{\sigma_1} - \frac{1}{2}\left[\frac{H\Sigma_1 H' + R}{\sigma_1^2} - \frac{H\Sigma_o H' + R}{\sigma_o^2}\right]$$

- Etapes du programme de calcul de T.

On calcule d'abord les facteurs correspondant aux deux modèles considérés : (o) et (1)

Ces facteurs sont : $(i = o,1)$

* $P_i = E\,X_i\,X_i^T$, P_i vérifie : $F_i\,P_i + P_i\,F'_i = -\,Q$

* K_i = gain de Kalman (correspondant au régime stationnaire)

K_i est donné par :

$$\begin{cases} \Sigma_i(t_o) = P_i \\ \text{entre deux observations (durée} = \Delta T) \\ \Sigma_i(t) = F_i\,\Sigma_i(t) + \Sigma_i(t)\,F'_i + Q \end{cases}$$

à l'instant d'une observation

$$\Sigma_i(t_+) = \Sigma(t_{i-}) - \Sigma(t_{i-})H'(H\Sigma(t_{i-})H' + R)^{-1} \cdot H \cdot \Sigma(t_{i-})$$

Soit $\Sigma_{i\infty}$ la solution stable correspond à ce système d'équation

$$K_i = \Sigma_{i\infty} H' \left(H\Sigma_{i\infty} H' + R\right)^{-1}$$

* $\sigma_i^2 = E\{y^2(t_{n-1})/y(t_o)\ldots y(t_{n-1})\}$ σ_i^2 est donné par : $\sigma_i^2 = H\,\Sigma_{i\infty}\,H' + R$

Ces facteurs sont indépendants du paramètre du vrai modèle (F)

Connaissant ces facteurs, et le paramètre du vrai modèle :
$F = F_o + \zeta\,(F_1 - F_o)$, on calcule les matrices Σ_1 et Σ_o d'après les équations exposées en appendice [III].
$E_\zeta\ \theta$ est alors donné par (6-6) d'où la valeur de T.

Trois programmes de calcul de T ont été faits :
Le premier considère F_o et F_1 fixes ; il calcule $T = f(F, R, \Delta T)$. Le second considère F_o et F_1 variable ; on fait varier R et ΔT, mais on prend $F = F_1$.

Dans le troisième programme, R et ΔT sont fixes, F_o est fixé également ; seuls F_1 et F sont variables. Ces trois programmes sont similaires, ils ne diffèrent que par la variation de tel ou tel facteur. Les organigrammes sont présentés en annexe [4].

- méthode numérique utilisée.
Le calcul de P, P_o, P_1 s'obtient par intégration d'une équation différentielle du type :

$$\dot{P} = FP + PF' + Q$$

Pour le calcul de K_o, K_1, et Σ_o, Σ_1, nous faisons intervenir des équations différentielles du même type.
Pour l'intégration de ces équations différentielles, nous avons utilisé la méthode de Runge et Kutta.

Cette méthode est la suivante :

Soit P_n.

On pose :

$$K_1 = \rho(FP_n + P_nF' + Q)$$

$$K_2 = K_1 + \frac{\rho}{2}(FK_1 + K_1F')$$

$$K_3 = K_1 + \frac{\rho}{2}(FK_2 + K_2F')$$

$$K_4 = K_1 + \rho(FK_3 + K_3F')$$

alors : $P_{n+1} = P_n + \frac{1}{6}(K_1 + 2K_2 + 2K_3 + K_4)$

Pour le calcul de P_o, P_1 et P, nous avons pris : $\rho = 0.0003$

Pour le calcul de K_o, K_1 et Σ_o, Σ_1, ρ varie de 0.0002 à 0.0016, le nombre d'itérations étant fixé à 10.

La période d'échantillonnage est telle que :

$$\Delta T = 10.\ \rho$$

3°) variation des facteurs.

Comme indiqué plus haut, nous avons fait varier la covariance du bruit de mesure R et la période d'échantillonnage ΔT.

Pour le système du premier ordre :

R varie de 0.0035 à 0.028

ΔT varie de 2 msec à 16 msec.

$F_o = \alpha_o = -50$

$F_1 = \alpha_1 = -45$

$F = \lambda$ varie de - 80 à - 5 pour chaque valeur de R et ΔT.

Pour le système du second ordre :

R varie de 0.0035 à 0.028

ΔT varie de 2 msec à 16 msec.

$$F_o = \begin{bmatrix} -50 & 2 \\ 0 & -1 \end{bmatrix}$$

$$F_1 = \begin{bmatrix} -45 & 2 \\ 0 & -1 \end{bmatrix}$$

$F = F_o + \varepsilon(F_1 - F_o)$

ζ varie de - 6 à + 7.

VI. 2. 4. Résultats.

Comme précisé au dessus, le temps d'observation T n'a pas été calculé quand $F \in]F_o, F_1[$, le résultat étant peu significatif. En annexe [5], nous présentons les différentes courbes donnant la valeur de T en fonction de F, et des différents facteurs du modèle.
Ces courbes sont divisées en deux groupes :

1°) F_o, F_1 fixes.

Pour chaque valeur de R, nous avons calculé les courbes $T = f(F_1, F_o, F, R, \Delta T)$.
Quatre valeurs de ΔT ont été prises en compte.
Pour une même valeur de ΔT et de R_1, plus F s'éloigne de (F_o, F_1) plus T devient petit. Ce résultat avait été montré dans le chapitre IV (théorème 4-4). De plus, on s'aperçoit que pour des valeurs de R élevées, l'écart entre les temps d'observation pour un même F et pour des périodes d'échantillonnage différentes est important.
Pour des valeurs faibles de R, ce même écart devient très faible.

Selon l'importance du bruit de mesure, le choix de la période d'échantillonnage pour obtenir un temps d'observation relativement faible est donc plus ou moins important. Plus R est élevée, plus ce choix s'impose.

2°) F_o fixe, F_1 variable et F variable.

R et ΔT sont fixes : R = 0.007, ΔT = 10 msec.
Pour chaque F_1, on calcule le temps d'observation en fonction de F.
Nous avons vu dans le premier groupe de courbes que T décroit quand F s'éloigne de F_1. On s'aperçoit ici que pour un même F, plus F_1 se rapproche de F, plus le temps d'observation devient petit.
Ici se pose le problème du choix de F_1 pour l'algorithme de détection de variation :

Si F_1 est éloigné de F_o, le temps d'observation sera faible, mais alors on détectera mal les valeurs de F telles que $F_o < F < F_1$. Il faut choisir la valeur de F_1 la plus éloignée de F_o, telle que si $F_o < F < F_1$ la détection de variation de paramètre n'a que peu d'importance.

VI. 3. Conclusion.

La simulation en temps réel des deux algorithmes nous a permis de mettre en évidence le temps de calcul et l'encombrement mémoire relativement faible, nécessaire à la mise en oeuvre.
Les courbes de visualisation montrent le comportement complexe de θ_n. La dispersion du nombre d'observations nécessaires pour détecter est relativement grande.

Le critère de performance étudié, T, est une expression approchée ou exacte à un facteur près, des temps moyens d'observation correspondant aux algorithmes proposés.
Il s'avère, qu'avec un bruit de mesure de covariance élevée, le choix de la période d'échantillonnage est important, vis à vis de la rapidité de la détection.
Enfin, pour l'algorithme de détection de variation, le choix du paramètre a_1 relève d'un compromis entre la rapidité et la sensibilité de détection de variation.

Chapitre VII

Conclusion Générale

L'étude du test de détection entre deux structures gaussiennes markoviennes nous a permis d'aborder le problème de la détection de changement dans les caractéristiques d'un processus à représentation gaussienne markovienne, ce changement n'étant pas connu exactement.

Le problème de la détection de changement par saut ou changement brusque (changement connu) a déjà fait l'objet de publications [9], [10]. L'algorithme présenté au chapitre III reprend certaines idées de Newbold et Ho [10]. Le chapitre IV étudie le comportement du test entre deux structures gaussiennes markoviennes dans le cas général. La convergence en temps fini du test est démontrée.
De plus, une expression simple du temps moyen d'observation nécessaire pour prendre une décision est donnée.
La fonction L(a) est difficile à établir exactement, les calculs étant très complexes.

Le problème de détecter un changement de paramètre quand la valeur du changement n'est pas connue exactement revêt une importance pratique considérable. Deux aspects différents de ce problème ont été envisagés :

- détection d'un changement brusque d'une part (on cherche à détecter le plus rapidement possible le moment τ du changement)
- détection de variation d'autre part (on teste à un moment donné si le paramètre a varié de façon sensible ou non).

L'étude se ramène dans les deux cas à des tests d'hypothèses "multiples" (plusieurs modèles sont envisagés).
On a présenté ici des algorithmes optimaux, caractérisés par leur simplicité et leur facilité de traitement numérique.

Ces algorithmes ont été simulés sur une installation hybride (calculateur analogique couplé à un calculateur numérique).

Ils sont très bien adaptés au temps réel : la détection est séquentielle, le temps de calcul et le volume de mémoire sont relativement faibles. Nous avons également fait une étude du critère de performance de ces algorithmes (variation du temps moyen d'observation en fonction des différents paramètres du système).

Dans ce travail, nous avons fait les hypothèses suivantes sur la nature du changement :

- une seule des caractéristiques du modèle de $y(t)$ est susceptible de changer.
- la valeur finale après changement n'est pas connue, mais la "direction" du changement est connue.

Il serait intéressant d'étudier le cas plus général où plusieurs paramètres sont susceptibles de varier et où la "direction du changement n'est pas connue.
Il faut alors définir une norme sur le modèle de $y(t)$.

Bibliographie.

[1] M.S. BARTLETT, An introduction to stochastic Processes. London : Cambridge University Press, 1962.

[2] J.J. BUSSGANG, D. MIDDELTON, Optimum Sequential Detection of Signals in Noise. IRE Trans. on IT pp. 5-18 August 1955.

[3] P. FAURRE, P. CHATAIGNIER, Identification en temps réel et en temps différé par factorisation de matrices de Hankel. Colloque Franco-suédois : conduite des procédés, IRIA (Octobre 1971).

[4] P. FAURRE, Navigation inertielle optimale et filtrage statistique, Dunod, 1971.

[5] P. FAURRE, Réalisations markoviennes de processus stationnaires (Doctorat Es Sciences, Déc. 1972)

[6] J.C. HANCOCK, P.A. WINTZ, Signal detection theory, New-York, Mac Graw Hill 1966.

[7] T. KAILATH, An innovation approach to least square estimation, pt. I : Linear Filtering in additive white noise, IEEE Trans. Automatic Control, Vol. AC-13, pp. 646-655, December 1968.

[8] R.E. KALMAN , R.S. BUCY, New results in Linear Filtering and Prediction Theory Journal of Basic Engineering, pp. 35-45, March 1960.

[9] V.Ya. LUMELSKII, Algorithm for detecting the time of change of properties of a random process, Moscow, Automatika i Telemekhanika n° 10, pp. 67-73 October 1972.

[10] P.M. NEWBOLD, Y.C. HO, Detection of changes in the Characteristics of a Gauss-Markov Process, IEEE Trans. on aerospace and electronic systems, Sept. 1968.

[11] E. PARZEN, An approach to time series Analysis, Ann. Math. Statist. vol 32, N° 4, December 1961.

[12] F.C. SCHWEPPE, Evaluation of likchood functions for gaussian signals IEEE Trans. Information Theory, vol. I.T.-11, pp. 61-70. January 1965.

[13] C. STEIN, A note on Cumulative Sums, Annals of Mathematical Statistics, Vol. 17 (1946) pp. 498-499.

[14] SHIRJAEV, Detection of malfunctions of a production process, Troriya Veroyatnostei i E[illegible] Primeneniya 8 N° 3 (1963).

[15] A. WALD, Sequential analysis, New-York, John Wiley and Sons, Inc. (1947).

[16] A. WALD, J. WOLFOWITZ, Optimum Character of the Sequential Probability Ratio Test. Annals of Mathematical Statistics. Vol. 19 p. 326 (1948).

[17] Bibliothèque de programmes T2000, Département Automatique Appliquée, IRIA - LABORIA - ROCQUENCOURT, France.

Appendice I

Calcul de la fonction L(a) à partir du test d'hypothèses H_0 , H_1 .

Hypothèse : Le dépassement de $\Lambda(\mathcal{Y}_n)$ au-dessus des bornes A ou B est supposé négligeable.

Posons : $\theta_n = \text{Log}\,[\Lambda(\mathcal{Y}_n)]$

$E\, e^{h\,\theta_n}$ s'écrit aussi : $E\, e^{h\,\theta_n} = \int_{\Omega} \left[\frac{p[\mathcal{Y}_n/H_1]}{p[\mathcal{Y}_n/H_0]}\right]^h p[\mathcal{Y}_n/H]\, d\mathcal{Y}_n$ (I-1)

Nous énonçons d'abord un résultat sur $E\, e^{h\,\theta_n}$

Théorème I-1. Il existe un h_0 et un seul vérifiant : $E\, e^{h_0 \theta_n} = 1$ $\quad h_0 \neq 0$

Démonstration :

Soit $g(h) = E\, e^{h\,\theta_n}$

1 - montrons tout d'abord que $\lim_{h \to \pm\infty} g(h) = +\infty$ (I-2)

θ_n étant une variable aléatoire, il existe $\delta > 0$ tel que : $p[e^{\theta_n} > 1] = \delta$
et il existe $\eta > 0$ tel que : $p[e^{\theta_n} < -1] = \eta$.

alors, pour $h \to +\infty$, on a : $g(h) > \delta\, e^{h}$ donc $g(h) \to +\infty$

pour $h \to -\infty$, on a : $g(h) > \eta\, e^{-h}$ et $g(h) \to +\infty$

2 - montrons maintenant que $\frac{d^2 g(h)}{d\,h^2} > 0$ (I-3)

Pour n'importe quelle valeur de h, g(h) existe. En différentiant, on obtient :

$$\frac{d^r g(h)}{d\,h^r} = E[\theta_n^r\ e^{h\,\theta_n}]$$

or $\theta_n^2\ e^{h\,\theta_n} > 0$, donc $\frac{d^2 g(h)}{d\,h^2} > 0$.

3 - la fonction g(h) est une fonction convexe (résultat (I-3)). Ses branches extrêmes tendent vers $+\infty$ pour $h \to \pm\infty$ (résultat (I-2)).
Il existe donc une seule valeur $h^* \neq 0$ telle que g(h) prenne sa valeur minimum pour $h = h^*$.

4 - nous pouvons maintenant terminer la démonstration du théorème :

On a : $\left[\frac{d\,g(h)}{dh}\right]_{h=o} = E\,\theta_n$ supposons $E\,\theta_n \neq o$, alors $h^* \neq o$ et,

De par la définition de h^* on a : $g(h^*) < g(o) = 1$.

Il est clair que $g(h)$ est monotone décroissante sur $]-\infty, h^*]$ et monotone croissante sur $[h^*, +\infty[$.

Comme $g(o) = 1$ et $g(h^*) < 1$, il existe une et une seule valeur $h_o \neq o$ telle que $g(h_o) = 1$

■

Remarque :

Si $E\,\theta_n < o$, alors $h^* > o$ et si $E\,\theta_n > o \rightarrow h^* < o$

donc h_o et $E\,\theta_n$ sont de signes contraires.

Posons $g(\mathcal{Y}_n/a) = [\Lambda(\mathcal{Y}_n)]^{h_o} p[\mathcal{Y}_n/a]$

Il est clair d'après le théorème I-1 que $g(\mathcal{Y}_n/a)$ est une densité de probabilité :

En effet : $\int_\Omega g(\mathcal{Y}_n/a)\,d\mathcal{Y}_n = E\,e^{h_o \theta_n} = 1$

Compte tenu de tout ces résultats, nous allons construire un nouveau test d'hypothèses avec $g(\mathcal{Y}_n/a)$ et $p(\mathcal{Y}_n/a)$ comme densités de probabilité, et A^{h_o}, B^{h_o} comme bornes.

Nous montrerons que ce nouveau test est équivalent au test des hypothèses H_o, H_1.

Considérons le nouveau test d'hypothèses :

H_g, la vraie densité de probabilité de $\mathcal{Y}_n$ est $g(\mathcal{Y}_n/a)$

H_p, " " " " " " $p(\mathcal{Y}_n/a)$

Les bornes sont A^{h_o} et B^{h_o}.

A ce test on associe les probabilités d'erreur α' et β' telles que :

$$A^{h_o} = \frac{1-\beta'}{\alpha'} \quad , \quad B^{h_o} = \frac{\beta'}{1-\alpha'}$$

Lemme I-1

Le test d'hypothèses H_g , H_p est équivalent au test d'hypothèses H_o, H_1.

Démonstration.

on a : $$\frac{g(\mathcal{Y}_n/a)}{p(\mathcal{Y}_n/a)} = \Lambda(\mathcal{Y}_n)^{h_o}$$

alors : $$\frac{g(\mathcal{Y}_n/a)}{p(\mathcal{Y}_n/a)} \leqslant B^h \Leftrightarrow \Lambda(\mathcal{Y}_n) \leqslant B$$

et : $$\frac{g(\mathcal{Y}_n/a)}{p(\mathcal{Y}_n/a)} \geqslant A^h \Leftrightarrow \Lambda(\mathcal{Y}_n) \geqslant A$$

Si le test (H_g, H_p) consuit à l'acceptation de H_p, le test (H_o, H_1) conduit à l'acceptation de H_o.

■

Du lemme I-1, on déduit le théorème suivant :

Théorème I-2

La probabilité d'accepter H_o quand a est vrai(c'est-à-dire L(a)) est la même que la probabilité d'accepter H_p quand $p[\mathcal{Y}_n/a]$ est la vraie distribution de $\mathcal{Y}_n$. A cette dernière probabilité est $1 - \alpha'$

D'où on déduit :

$$(I\text{-}4) \qquad L(a) = 1 - \alpha' = \frac{A^{h_o} - 1}{R^{h_o} - B^{h_o}}$$

où h_o est l'unique solution $\neq o$ de $E\, e^{h\,\theta_n} = 1$.

Remarques :

* Compte tenu de la relation (I-1), du théorème (I-1) et du résultat (I-4), on peut facilement remarquer que :

$$h_o(a = a_o) = 1$$
$$h_o(a = a_1) = -1$$

D'où $L(a_o) = 1 - \alpha$

$L(a_1) = \beta$

* On s'aperçoit que pour $h_o = o$, il y a <u>indétermination</u> dans la relation (I-4). Soit $a = a'$, la valeur du paramètre correspondant à $h_o = o$.

On applique la règle de l'Hospital à (I-4) :
Au point a' on obtient :

$$L(a') = \lim_{h \to o} \frac{d/dh.(A^h - 1)}{d/dh.(A^h - B^h)}$$

$$= \lim_{h \to o} \frac{A^h \text{ Log } A}{A^h \text{ Log } A - B^h \text{ Log } B}$$

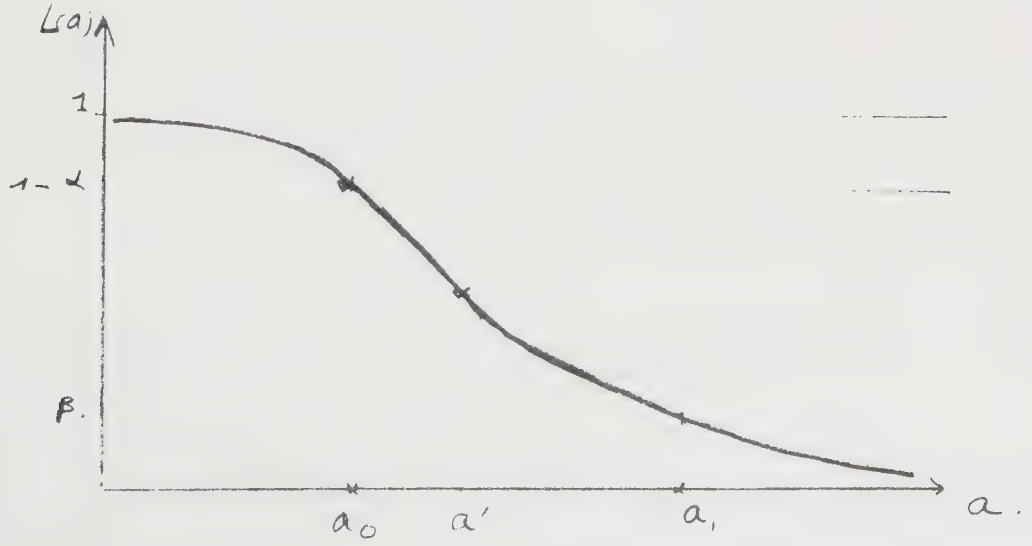
d'où :

(I-5) $$L(a') = \frac{\text{Log } A}{\text{Log } A/B}$$

Dans le cas où $\alpha = \beta$, on a : $A = 1/B$ et :

$$L(a') = 1/2.$$

* <u>La représentation d'une fonction L(a) typique</u> est donnée par la figure suivante (voir Bussgang et Middleton [2]).



Appendice II

Equations du filtre de Kalman-Bucy :

- modèle continu, observations discrètes.
- modèle discret, observations discrètes.

Cette appendice a pour seul but de rappeler les équations récursives du filtre récursif de Kalman-Bucy dans le cas d'observations discrètes. Pour plus de détails, on pourra consulter Kalman [8], ou Faurre [4].

1 - Modèle continu, observations discrètes.

Etant donné un processus markovien x(t) tel que :

$$\dot{x}(t) = F\,x(t) + v(t) \quad \text{avec,} \quad E\{v(t).v(\tau)\} = Q.\delta(t-\tau)$$

et les observations $y(t_i)$ fournies à une suite d'instants discrets.

$$y(t_i) = H\,x(t_i) + w(t_i) \quad \text{avec,} \quad E\{w(t_i).w(t_j)\} = R\,\delta_{ij}$$

on cherche à calculer par un algorithme récursif la meilleure estimée de x(t) :

$$(II\text{-}1) \qquad \hat{x}(t) = E\{x(t) / y(t_0), y(t_1), \ldots y(t_i)\} \quad , \; t_i \leq t < t_{i+1}$$

Soit : $\Sigma(t) = E\{(x-\hat{x}(t))(x-\hat{x}(t))'\}$.

Les équations récursives permettant de calculer $\hat{x}(t)$ sont les suivantes :

$$(II\text{-}2) \quad \begin{cases} \text{- Entre deux observations :} \\ \quad t_{i-1} < t < t_i \\ \qquad \dot{\hat{x}}(t) = F\hat{x}(t) \\ \qquad \dot{\Sigma}(t) = F\Sigma(t) + \Sigma(t)F' + Q \\ \text{— à l'instant d'une observation } t = t_i \\ \qquad \hat{x}(t_{i+}) = \hat{x}(t_{i-}) + K(t_i)\left[y(t_i) - H\hat{x}(t_{i-})\right] \\ \qquad K(t_i) = \Sigma(t_{i-})H'\left[H\Sigma(t_{i-})H' + R\right]^{-1} \\ \qquad \Sigma(t_{i+}) = \Sigma(t_{i-}) - \Sigma(t_{i-})H'\left[H\Sigma(t_{i-})H' + R\right]^{-1}H\Sigma(t_{i-}) \\ \text{et} \qquad \sigma^2(t_i) = H\Sigma(t_{i-})H' + R \end{cases}$$

Remarque

Au bout d'un certain temps la matrice $\Sigma(t)$ tend vers une solution d'équilibre :

$$\Sigma(t) = \Sigma_\infty$$

Alors : $K(t) = K_\infty$

$$\sigma^2(t) = \sigma^2_\infty$$

Le filtre est dit en régime stationnaire.

Les équations donnant $\underline{y(t_n)}$ et $\sigma^2(t_n)$ sont alors les suivantes :

$$E\{y(t_n)/Y_{n-1}\} = H\, E\{x(t_n)/Y_{n-1}\}$$

$$\mathrm{cov}\{y(t_n)/Y_{n-1}\} = H\, \mathrm{cov}\{x(t_n)/Y_{n-1}\}\, H' + R$$

c'est-à-dire :

$$\text{(II-3)} \qquad \begin{cases} \hat{y}(t_n) = H\, \hat{x}(t_n -) \\ \sigma^2(t_n) = H\, \Sigma(t_{n-1})\, H' + R \end{cases}$$

2 - Modèle discret, observations discrètes.

Nous employons les mêmes notations que dans le paragraphe précédent.

Le modèle utilisé est ici le suivant : $x(t_{n+1}) = \phi\, x(t_n) + v(t_n)$

$$y(t_n) = H\, x(t_n) + w(t_n)$$

On cherche à calculer : $\hat{x}(t_n) = E\{x(t_n) / y(t_0), y(t_1), \ldots y(t_n)\}$

Les équations donnant $\hat{x}(t_n)$ sont les suivantes :

$$\text{(II-4)} \qquad \begin{cases} \hat{x}(t_{n+1}) = \phi\, \hat{x}(t_n) + \phi\, K(t_n)\, [y(t_n) - H\, \hat{x}(t_n)] \\ K(t_n) = P(t_n)\, H' R^{-1} \\ M(t_n) = \phi\, P(t_n)\, \phi' + Q \\ P(t_n) = M(t_n) - M(t_n)\, H' (H\, M(t_n)\, H' + R)^{-1} H\, M(t_n) \end{cases}$$

Remarque :

Comme dans le paragraphe précédent, $P(t_n)$ a une solution d'équilibre

$$P(t_n) = P_\infty$$

alors : $K(t_n) = K_\infty$

Le filtre est alors en régime stationnaire.

Les équations donnant $\hat{y}(t_n)$ et $\sigma^2(t_n)$ sont les suivantes :

$$\text{(II-5)} \qquad \begin{cases} \hat{y}(t_n) = H\, \hat{x}(t_n) \\ \sigma^2(t_n) = H\, M(t_n)\, H' + R \end{cases}$$

Appendice III

Calcul de $E\{(X(t)-\hat{X}_f(t)).(X(t)-\hat{X}_f(t))'\}$

On considère les deux modèles suivants :

- le modèle exact représentant le signal observé y(t) :

(1)
$$\begin{cases} \dot{x}(t) = F\,x(t) + v(t) \\ y(t_n) = H\,x(t_n) + w(t_n) \end{cases}$$
avec $E\,x(t_0)\,x'(t_0) = \Lambda_0$
$E\,v(t)\,v'(t) = Q\,\delta(t-\tau)$
$E\,w(t_i)\,w(t_j) = R\,\delta_{ij}$
(observations discrètes)

- le modèle utilisé servant à définir l'algorithme de filtrage :

(2)
$$\dot{x}_f(t) = F_f\,x_f(t) + v_f(t) \quad \text{avec} \quad E\,x_f(t_0)\,x_f'(t_0) = \Lambda_{0f}$$
$$y(t_n) = H\,x_f(t_n) + w(t_n) \qquad E\,v_f(t)\cdot v_f'(t) = Q_f\,\delta(t-\tau)$$
(observations discrètes)

Nous nous plaçons dans le cas général où le modèle utilisé ne correspond pas au modèle vrai.

On cherche alors l'expression de la vraie covariance d'erreur de filtrage :

$$\Sigma(t) = E\left\{\left(x(t)-\hat{x}_f(t)\right)\cdot\left(x(t)-\hat{x}_f(t)\right)'\right\}$$

On suppose le filtre en régime stationnaire.

En regroupant les équations (1) et celles du filtre (appendice II) on obtient :

- entre deux observations :
$$\dot{x}(t) = F\,x(t) + v(t)$$
$$\dot{\hat{x}}_f(t) = F\,\hat{x}_f(t)$$

- à l'instant d'une observation :
$$x(t_+) = x(t_-)$$
$$\hat{x}_f(t_+) = \hat{x}_f(t_-) + K_f\left[y(t) - H\hat{x}_f(t_-)\right]$$
$$\hat{x}_f(t_+) = (I - K_f H)\,\hat{x}_f(t_-) + K_f H x(t) + K_f w(t)$$

ou

Soit : $\Pi(t) = E\left\{\begin{bmatrix} x(t) \\ \hat{x}_f(t) \end{bmatrix}\begin{bmatrix} x'(t) & \hat{x}_f'(t) \end{bmatrix}\right\}$

Alors $\Pi(t)$ est obtenue par la condition initiale.

$$\Pi(t_0) = \begin{bmatrix} \Lambda_0 & 0 \\ 0 & 0 \end{bmatrix}$$

et par les équations :

- entre deux observations :

$$\dot{\Pi}(t) = \mathcal{F}\,\Pi(t) + \Pi(t)\,\mathcal{F}' + \mathcal{Q} \qquad \text{où} \qquad \mathcal{F} = \begin{bmatrix} F & 0 \\ 0 & F_f \end{bmatrix}$$
$$\mathcal{Q} = \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix}$$

- à l'instant d'une observation :

$$\Pi(t_+) = \begin{bmatrix} I & 0 \\ k_t H & I - k_t H \end{bmatrix} \Pi(t_-) \begin{bmatrix} I & 0 \\ k_t H & I - k_t H \end{bmatrix}' + \begin{bmatrix} 0 & 0 \\ 0 & k_t R k_t' \end{bmatrix}$$

Connaissant $\Pi(t)$, on en déduit $\Sigma(t)$ par la relation suivante :

$$\Sigma(t) = [I \; , -I] \, \Pi(t) \begin{bmatrix} I \\ -I \end{bmatrix}$$

<u>Remarque</u> : $\Pi(t)$ au bout d'un certain temps tend vers <u>une solution d'équilibre</u>

$$\Pi(t) = \Pi_\infty$$

Compte tenu du fait que le filtre est en régime stationnaire et que le régime transitoire pour $\Pi(t)$ est négligeable, nous prenons pour valeur de $\Sigma(t)$:

$$\Sigma(t) = \Sigma_\infty$$

ou

$$\Sigma = [I \; , -I] \, \Pi_\infty \begin{bmatrix} I \\ -I \end{bmatrix}$$

Appendice IV

Théorème sur les valeurs propres d'une matrice triangulaire inférieure par blocs

Théorème :

Soit F, une matrice triangulaire inférieure par blocs :

$$F = \begin{bmatrix} A & 0 \\ B & C \end{bmatrix}$$

alors, l'ensemble des valeurs propres de F est la concaténation de l'ensemble des valeurs propres de A et de celui des valeurs propres de C.

Démonstration :

Nous allons montrer que F, matrice triangulaire inférieure par blocs, peut être mise, par changement de base, sous forme de matrice triangulaire inférieure.

Soient deux matrices, T_{11} et T_{22} telles que :

$T_{11} A T_{11}^{-1} = A_d$ soit diagonale.

$T_{22} C T_{22}^{-1} = C_d$ soit diagonale.

Soient également : $T_1 = \begin{bmatrix} T_{11} & 0 \\ 0 & I \end{bmatrix}$,

$T_2 = \begin{bmatrix} I & 0 \\ 0 & T_{22} \end{bmatrix}$

alors on a : $T_1 F T_1^{-1} = \begin{bmatrix} A_d & 0 \\ B' & C \end{bmatrix}$

et : $T_2 T_1 F T_1^{-1} T_2^{-1} = \begin{bmatrix} A_d & 0 \\ B'' & C_d \end{bmatrix}$

en posant $T_3 = T_2 T_1$ on a toujours : $T_3 F T_3^{-1} = \begin{bmatrix} A_d & 0 \\ B'' & C_d \end{bmatrix}$

$T_3 F T_3^{-1}$ est triangulaire inférieure. Les valeurs propres de F sont sur la diagonale principale. Le théorème est démontré.

Appendice V

Théorème de monotonie croissante pour $E_\zeta\ \theta$ (F scalaire)

L'expression de $E_\zeta\ \theta$ est :

(V-1) $$E_\zeta \theta = \log \frac{\sigma_0}{\sigma_1} - \frac{1}{2}\left[\frac{C_{01}}{\sigma_1^2} - \frac{C_{00}}{\sigma_0^2}\right]$$

où $$C_{oi} = H \Sigma_{oi} H' + R \quad (i = 0, 1)$$

Le calcul de Σ_{oi} (explicité en appendice III) est donné par le système d'équations :

(V-2) $$\begin{cases} \text{- entre deux observations :} & \dot{\pi}_i(t) = \mathcal{F}_i \pi_i(t) + \pi_i(t) \mathcal{F}_i' + \mathcal{Q} \quad \text{avec :} \ \mathcal{F}_i = \begin{bmatrix} F & 0 \\ 0 & F_i \end{bmatrix}, \ \mathcal{Q} = \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix} \\ \text{- à l'instant d'une observation :} & \pi_i(t_+) = \begin{bmatrix} I & 0 \\ K_i H & I - K_i H \end{bmatrix} \pi_i(t_-) \begin{bmatrix} I & 0 \\ K_i H & I - K_i H \end{bmatrix}' + \begin{bmatrix} 0 & 0 \\ 0 & K_i R K_i' \end{bmatrix} \end{cases}$$

Soit $\pi_i = \pi_{i\infty}$

alors : $$\Sigma_{oi} = \begin{bmatrix} I & -I \end{bmatrix} \pi_i \begin{bmatrix} I \\ -I \end{bmatrix}$$

Nous allons nous servir du résultat du lemme suivant :

Lemme V-1

Soit l'équation différentielle :

(V-3) $$\dot{\pi}(t) = \begin{bmatrix} F & 0 \\ K_c H & F_i - K_c H \end{bmatrix} \pi(t) + \pi(t) \begin{bmatrix} F & 0 \\ K_c H & F_i - K_c H \end{bmatrix}' + \begin{bmatrix} Q & 0 \\ 0 & K_c R K_c' \end{bmatrix}$$

alors, si l'on pose : (V-4) $$K_c = K_c \cdot \Delta t \qquad R = R_c / \Delta t$$

les deux formes (V-2) et (V-3) génèrent la même covariance (au premier ordre en Δt).

Preuve :

On tire aisément des relations (V-2) et (V-4) :

$$\frac{\pi_i(t_{n+1}) - \pi_i(t_n)}{\Delta t} = \mathcal{F}_i \pi_i(t_n) + \pi_i(t_n) \mathcal{F}_i' + \begin{bmatrix} 0 & 0 \\ K_c H & -K_c H \end{bmatrix} \pi_i(t_n) + \pi_i(t_n) \begin{bmatrix} 0 & K_c H' \\ 0 & -K_c H' \end{bmatrix}$$

$$+ \begin{bmatrix} 0 & \cdot \\ 0 & K_c R K_c' \end{bmatrix} + \begin{bmatrix} Q & \cdot \\ 0 & \cdot \end{bmatrix} + \mathcal{O}(\Delta t)$$

c'est-à-dire :

$$\frac{\Pi_i(t_{n+1}) - \Pi_i(t_n)}{\Delta t} = \begin{bmatrix} F & 0 \\ K_{ic}H & F_i - K_{ic}H \end{bmatrix} \Pi_i(t_n) + \Pi_i(t_n) \begin{bmatrix} F & 0 \\ K_{ic}H & F_i - K_{ic}H \end{bmatrix}' + \begin{bmatrix} Q & 0 \\ 0 & K_{ic} R_c K_{ic}' \end{bmatrix} + \mathcal{O}(\Delta t)$$

Ce résultat est vrai quelque soit les paramètres susceptibles de varier.
De plus, nous savons que K_{fc} est la valeur d'équilibre du gain de Kalman correspond au modèle de paramètres : F, Q, H, R_c (modèle continu).

Le paramètre susceptible de changer est la constante de temps F scalaire.
Nous allons calculer explicitement la valeur de $E_\zeta\ \theta$ (au premier ordre).

Posons H = 1 pour simplifier les calculs (cela n'enlève rien à la généralité des résultats).

Soit $\Pi_i = \begin{bmatrix} a_i & b_i \\ b_i & c_i \end{bmatrix}$ la solution d'équilibre de l'équation (V-3).

On a : $$\begin{bmatrix} F & 0 \\ K_i & F_i - K_i \end{bmatrix} \begin{bmatrix} a_i & b_i \\ b_i & c_i \end{bmatrix} + \begin{bmatrix} a_i & b_i \\ b_i & c_i \end{bmatrix} \begin{bmatrix} F & 0 \\ K_i & F_i - K_i \end{bmatrix}' + \begin{bmatrix} Q & 0 \\ 0 & K_i^2 R \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

d'où l'on déduit :

$$a_i = \frac{Q}{2F}$$

$$b_i = \frac{Q K_i}{2F(F + F_i - K_i)}$$

$$c_i = \frac{-Q K_i^2}{2F(F + F_i - K_i)(F_i - K_i)} - \frac{K_i^2 R}{2(F_i - K_i)}$$

d'où l'on tire :

$$\Sigma_{oi} = \begin{bmatrix} 1 & -1 \end{bmatrix} \Pi_i \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\Sigma_{oi} = a_i - 2b_i + c_i$$

(V-4)
$$\Sigma_{oi} = -\frac{Q}{2F} + \frac{Q^2}{2R}\left[\frac{1}{F(F + F_i - K_i)(F_i - K_i)} - \frac{1}{(F + F_i - K_i)^2 (F_i - K_i)}\right]$$

Un premier résultat sur Σ_{oi} est le suivant :

Lemme V-2

Σ_{oi} est une fonction monotone croissante de $|\Delta F|$

Preuve :

Posons $\Delta F = F_i - F$, F_i variable et F fixe.

il suffit de montrer que $\frac{d\Sigma_{oi}}{d\Delta F_i} > 0 \ \forall |\Delta F| > a$ c'est-à-dire :

$$\frac{d\Sigma_{oi}}{dF_i} \begin{cases} > 0 & \text{si } \Delta F > 0 \\ < 0 & \text{si } \Delta F < 0. \end{cases}$$

Compte tenu de (V-4), on obtient :

$$\frac{d\Sigma_{oi}}{dF_i} = \frac{1}{(F_i - K_i)^2}\left[\frac{F_i + 2(F_i - K_i)}{(2F_i - K_i)^2} - \frac{F_i}{F}\frac{F + 2(F_i - K_i)}{(F + F_i - K_i)^2}\right]$$

alors :

si $F = F_i$ $\quad \frac{d\Sigma_{oi}}{dF_i} = 0.$

si $F_i < F$ ou $\Delta F < 0$, alors $\frac{d\Sigma_{oi}}{dF_i} < 0.$

si $F_i > F$, ou $\Delta F > 0$, alors $\frac{d\Sigma_{oi}}{dF_i} > 0.$

■

De ce lemme, on déduit le théorème suivant :

Théorème V-1

Soit deux modèles correspondant à F_o et F_1 et le vrai modèle correspond à F ; alors plus F s'éloigne de $[F_o, F_1]$, plus $|\Sigma_{01} - \Sigma_{00}|$ croît.

Démonstration :

Nous allons calculer $\frac{d[\Sigma_{01} - \Sigma_{00}]}{dF}$; sachant que :

$$\Sigma_{01} - \Sigma_{00} = \frac{\varphi^2}{2R}\left[\frac{1}{F(F_1 - K_1)(F + F_1 - K_1)} - \frac{1}{F(F_0 - K_0)(F + F_0 - K_1)} \frac{1}{(F_1 - K_1)(2F_1 - K_1)^2} + \frac{1}{(F_0 - K_0)(2F_0 - K_0)^2}\right]$$

alors : $$\frac{d[\Sigma_{01} - \Sigma_{00}]}{dF} = \frac{\varphi^2}{2R}\left[\frac{2F + F_0 - K_0}{(F_0 - K_0)(F + F_0 - K_0)^2} - \frac{2F + F_1 - K_1}{(F + F_1 - K_1)^2 (F + F_1 - K_1)^2}\right]$$

Choisissons $F_o < F_1$, alors :

$$\frac{2F + F_1 - K_1}{\quad} > \frac{2F + F_0 - K_0}{F_0 - K_0} \quad \text{et} \quad \frac{1}{(F + F_1 - K_1)^2} > \frac{1}{(F + F_0 - K_0)^2}$$

on voit alors que :

$$(V-5) \qquad \frac{d[\Sigma_{01} - \Sigma_{00}]}{dF} < 0$$

d'où : si $F < F_0 < F_1$, alors $\Sigma_{01} - \Sigma_{00} > 0$. (lemme V-2)

et, si $dF < 0$ F s'éloigne de $[F_0, F_1]$ et $\Sigma_{01} - \Sigma_{00}$ croît.

si $F_0 < F_1 < F$, alors $\Sigma_{01} - \Sigma_{00} > 0$ (lemme V-2)

et si $dF > 0$, F s'éloigne de $[F_0, F_1]$ et $\Sigma_{01} - \Sigma_{00}$ décroît.

Donc quand F s'éloigne de $[F_0, F_1]$, $|\Sigma_{01} - \Sigma_{00}|$ croît.

ou encore, $\Sigma_{01} - \Sigma_{00}$ est une fonction monotone décroissante de F.

□

(V-1) s'écrit aussi : $E_\zeta \theta = \log \frac{\sigma_0}{\sigma_1} - \frac{1}{2}\left[\frac{H\Sigma_{01}H' + R}{\sigma_1^2} - \frac{H\Sigma_{00}H' + R}{\sigma_0^2}\right]$

en tenant compte de l'approximation du premier ordre en Δt et du fait que H = 1 :

on a : $\sigma_0^2 \sim \sigma_1^2$ et $E_\zeta \theta = -\frac{1}{2}\left[\Sigma_{01} - \Sigma_{00}\right]$. (V-6).

De (V-6) et du théorème V-1, on tire le résultat suivant :

Théorème V-2 :

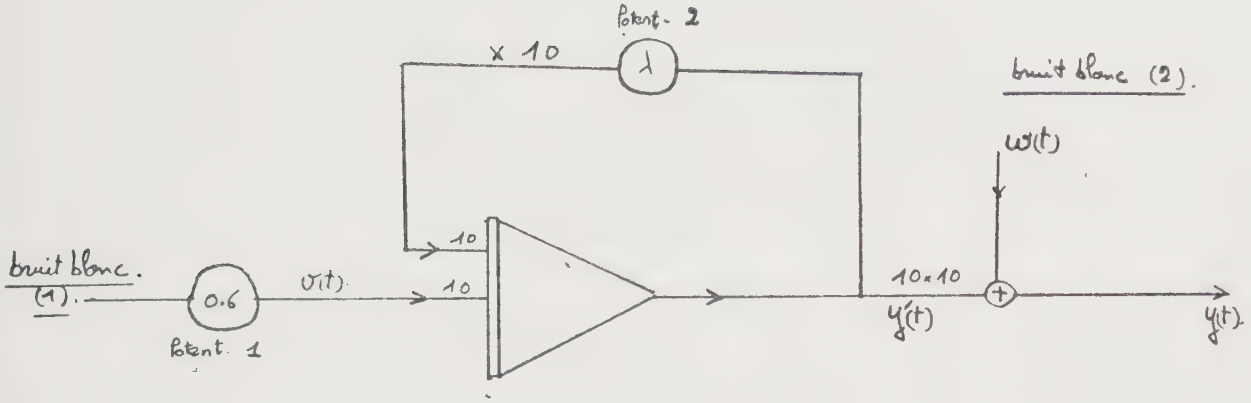
$E_\zeta \theta$ est une fonction monotone croissante de F, donc de ζ.

(puisque $F = (1-\zeta)F_0 + \zeta F_1$).

—

ANNEXE 1

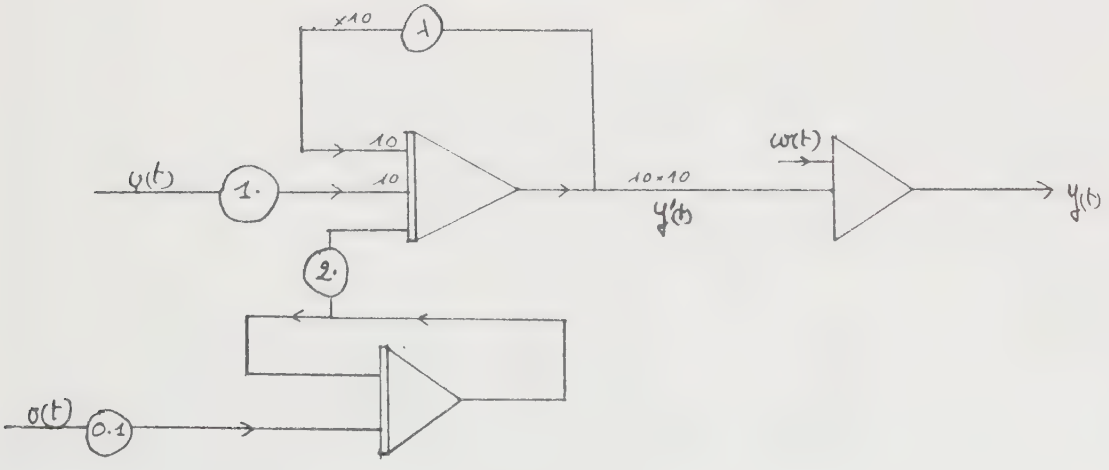
SCHEMA ANALOGIQUE SYSTEME D'ORDRE 1



BRUIT BLANC (1) : SEQUENCE INFINIE FREQUENCE DE COUPURE : 5 KHZ

BRUIT BLANC (2) : SEQUENCE INFINIE FREQUENCE DE COUPURE : 500 HZ

SCHEMA ANALOGIQUE SYSTEME D'ORDRE 2



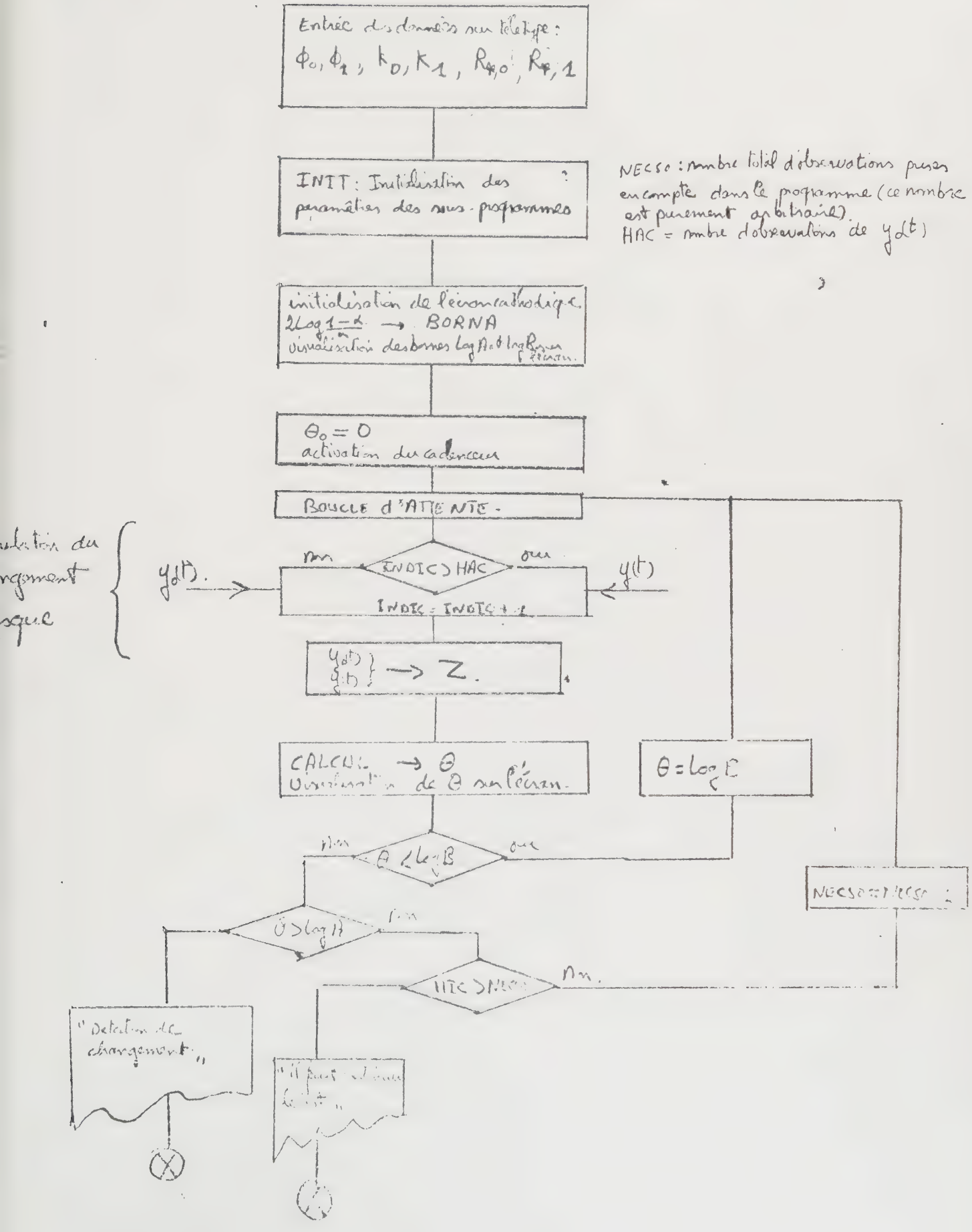
BRUIT BLANC V(t) : SEQUENCE INFINIE FREQUENCE DE COUPURE : 1500 HZ

BRUIT BLANC V'(t) : SEQUENCE INFINIE FREQUENCE DE COUPURE : 500 HZ

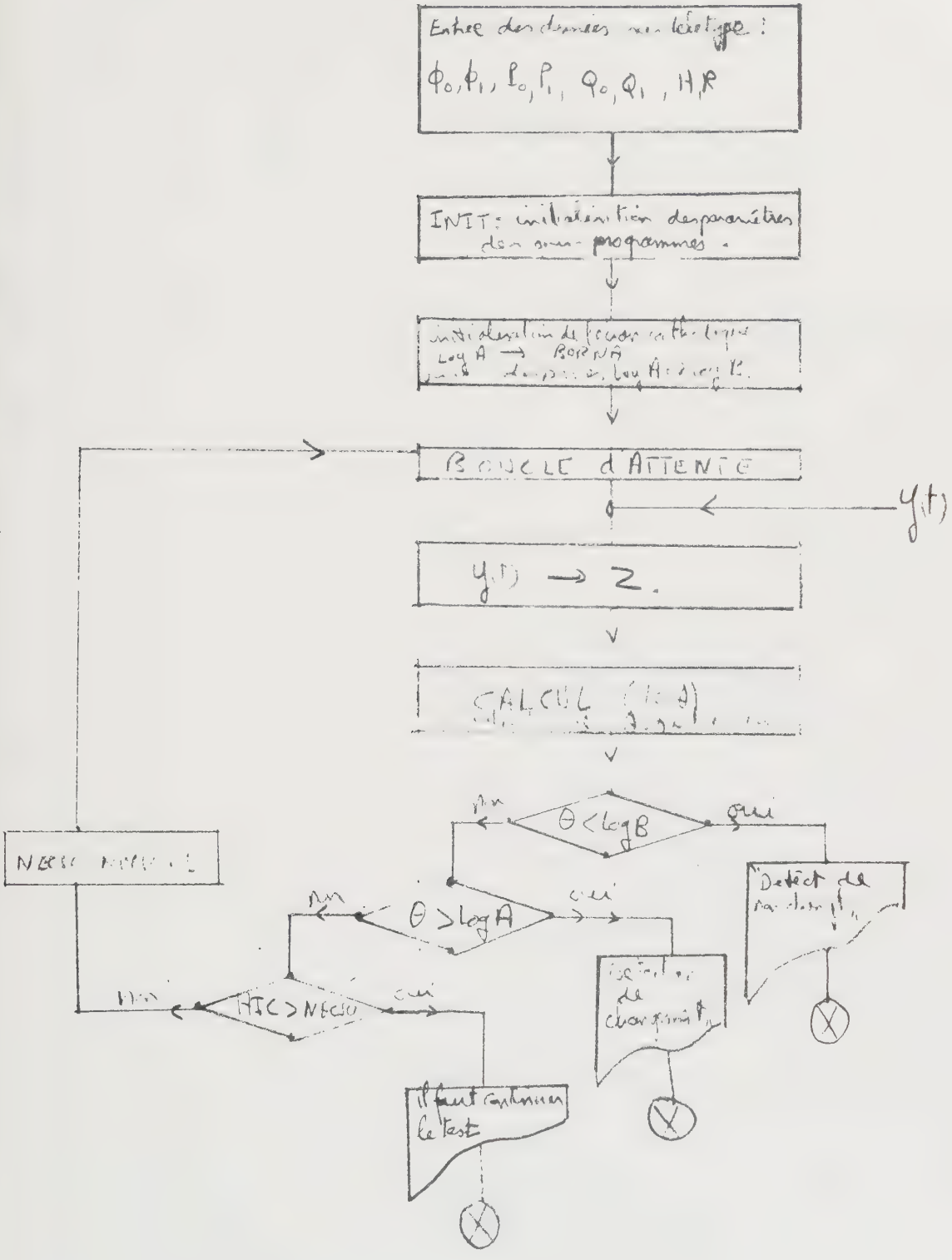
ANNEXE 2

SEC DECHT

Détection de changement brusque de paramètre

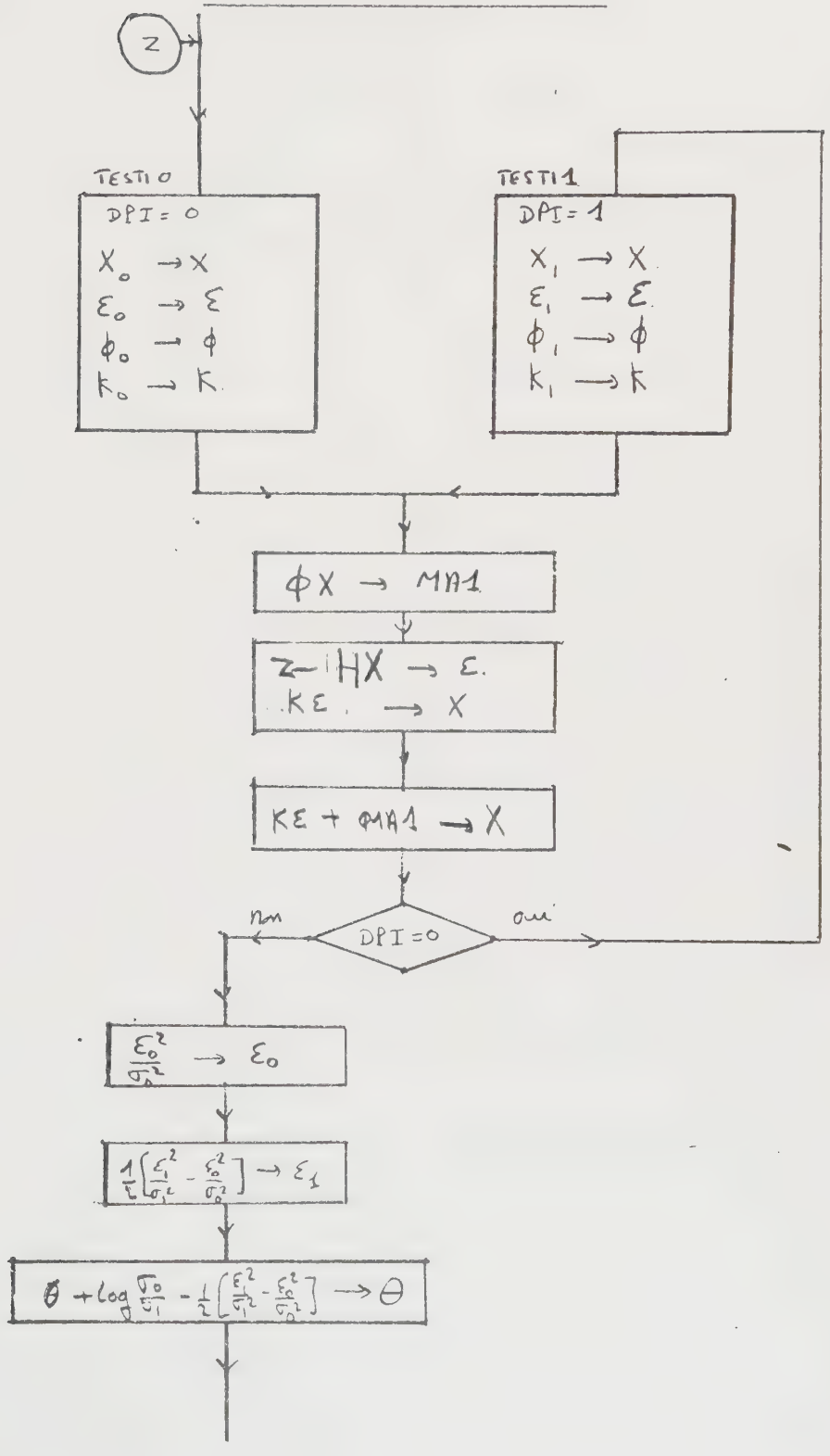


Détection de variation de paramêtre

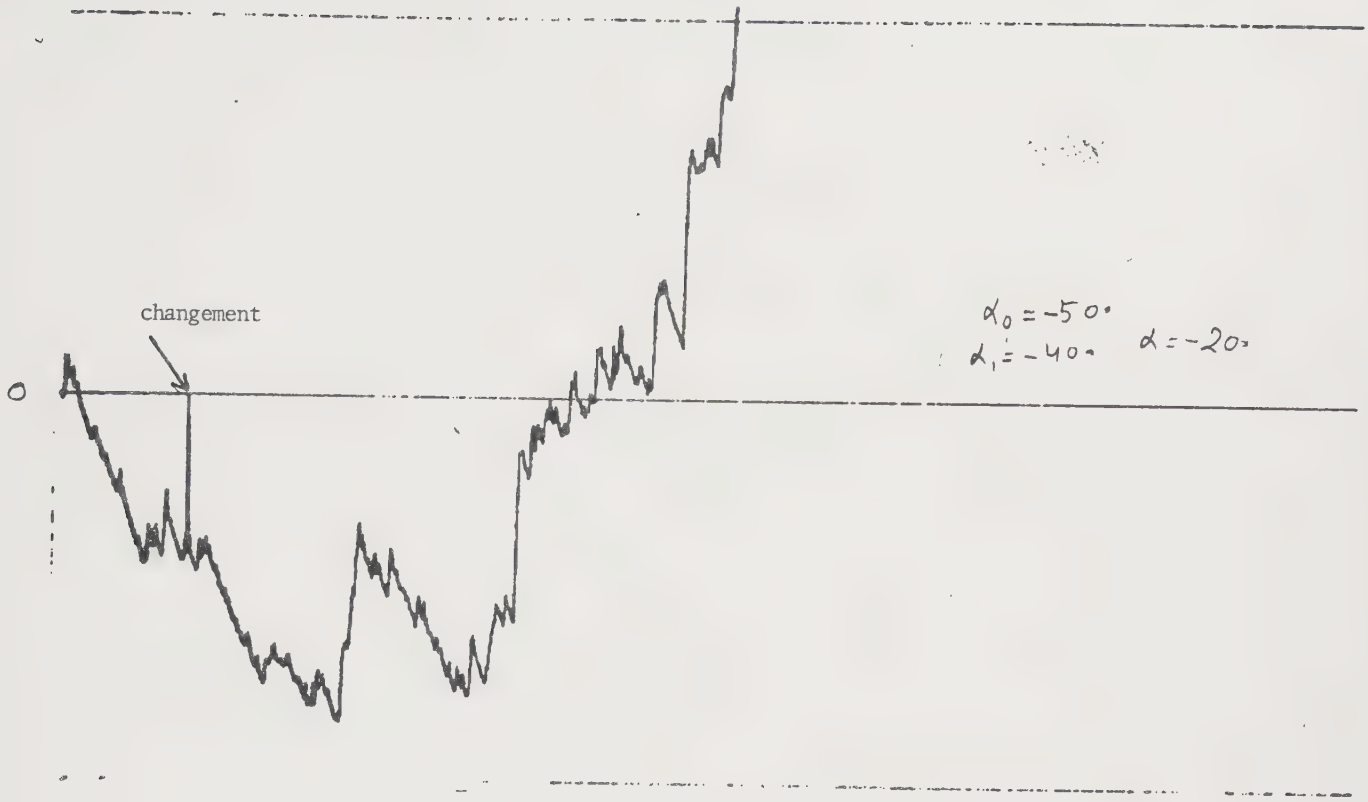


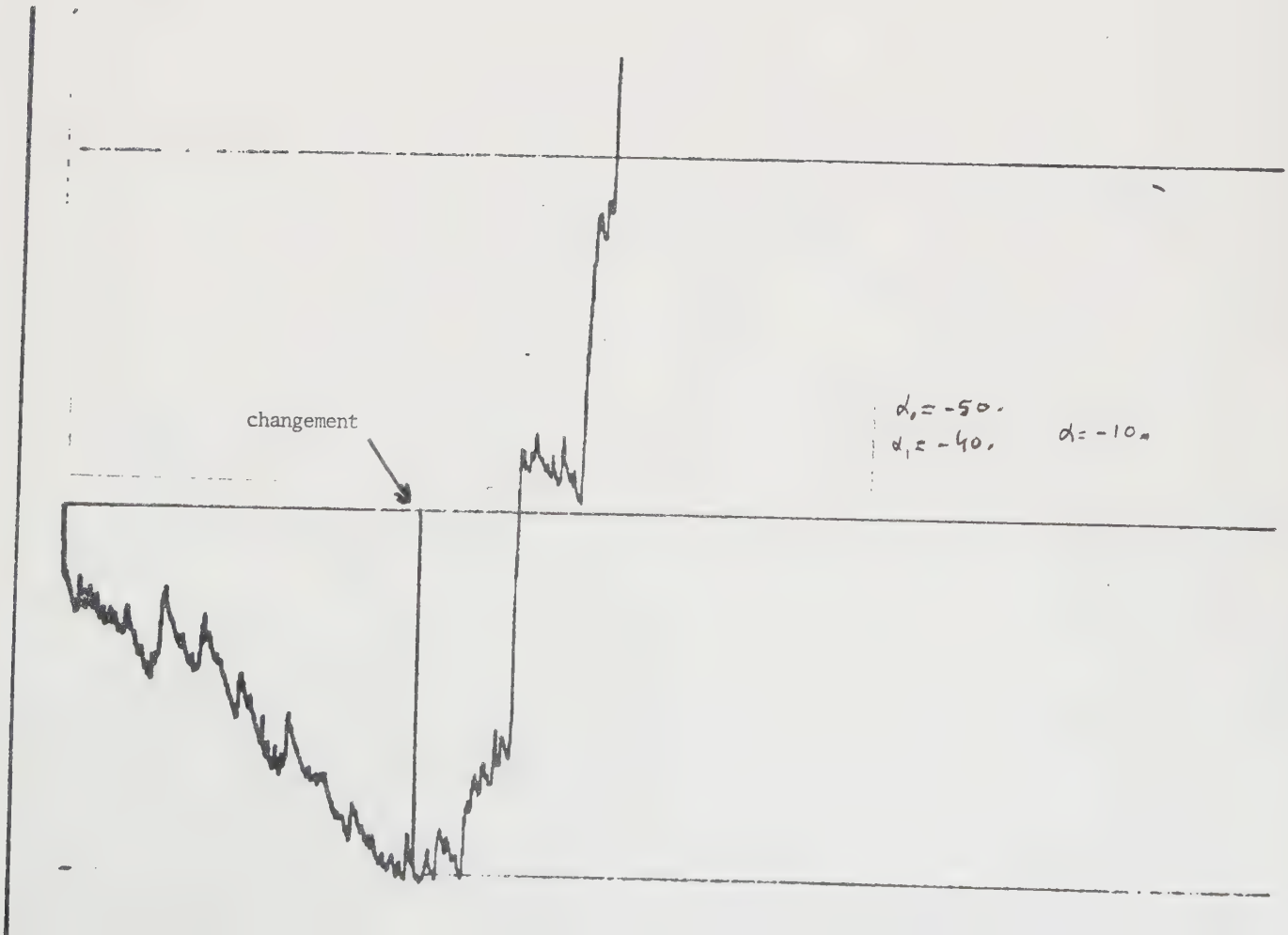
CALCUL

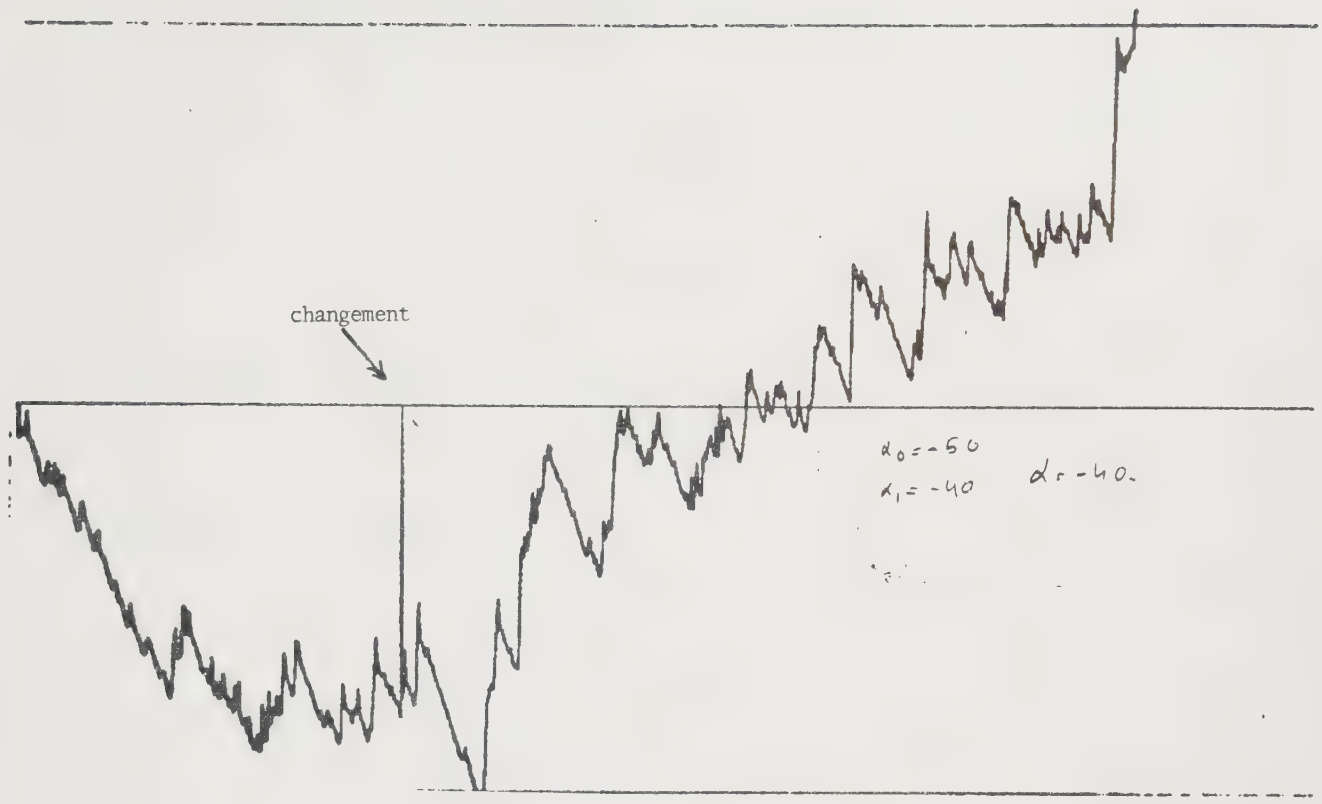
Calcul de θ_n connaissant $y(t_n)=Z$

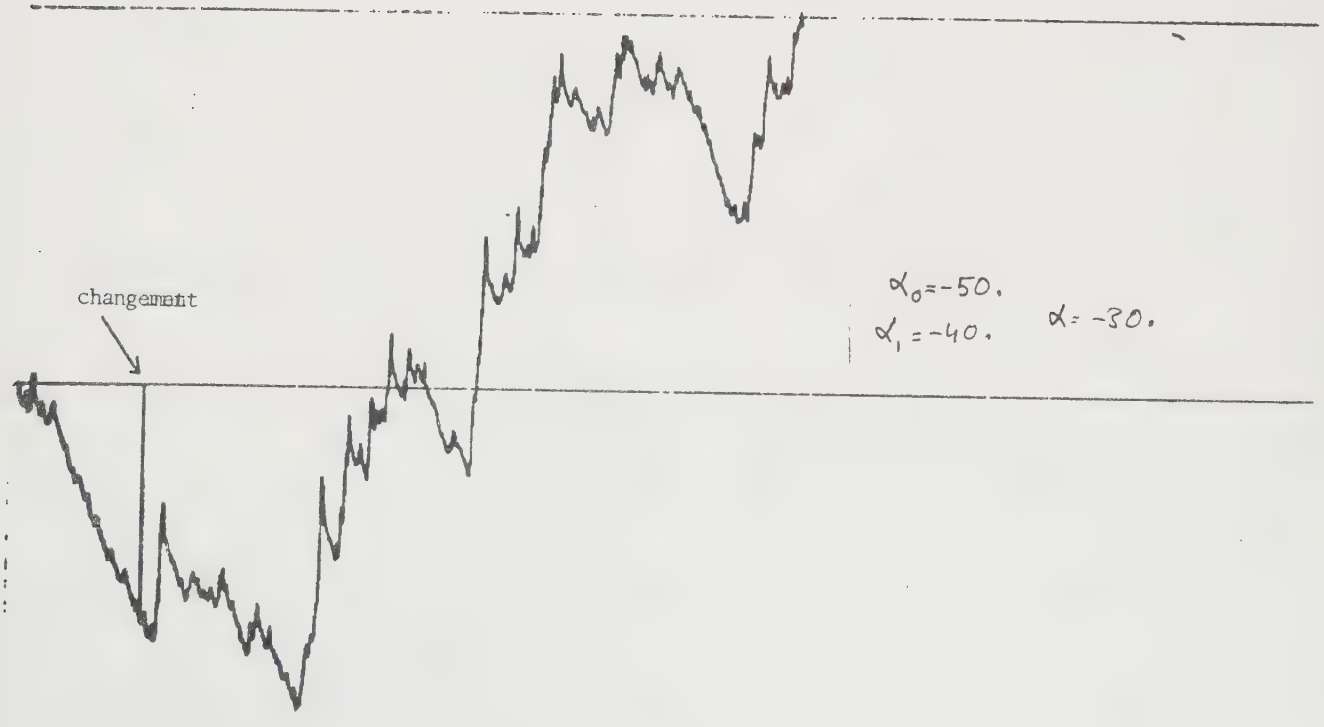


ANNEXE 3


changement
0
$\alpha_0 = -50°$
$\alpha_1 = -40°$
$\alpha = -20°$


changement
$\alpha_1 = -50°$
$\alpha_1 = -40°$
$\alpha = -10°$

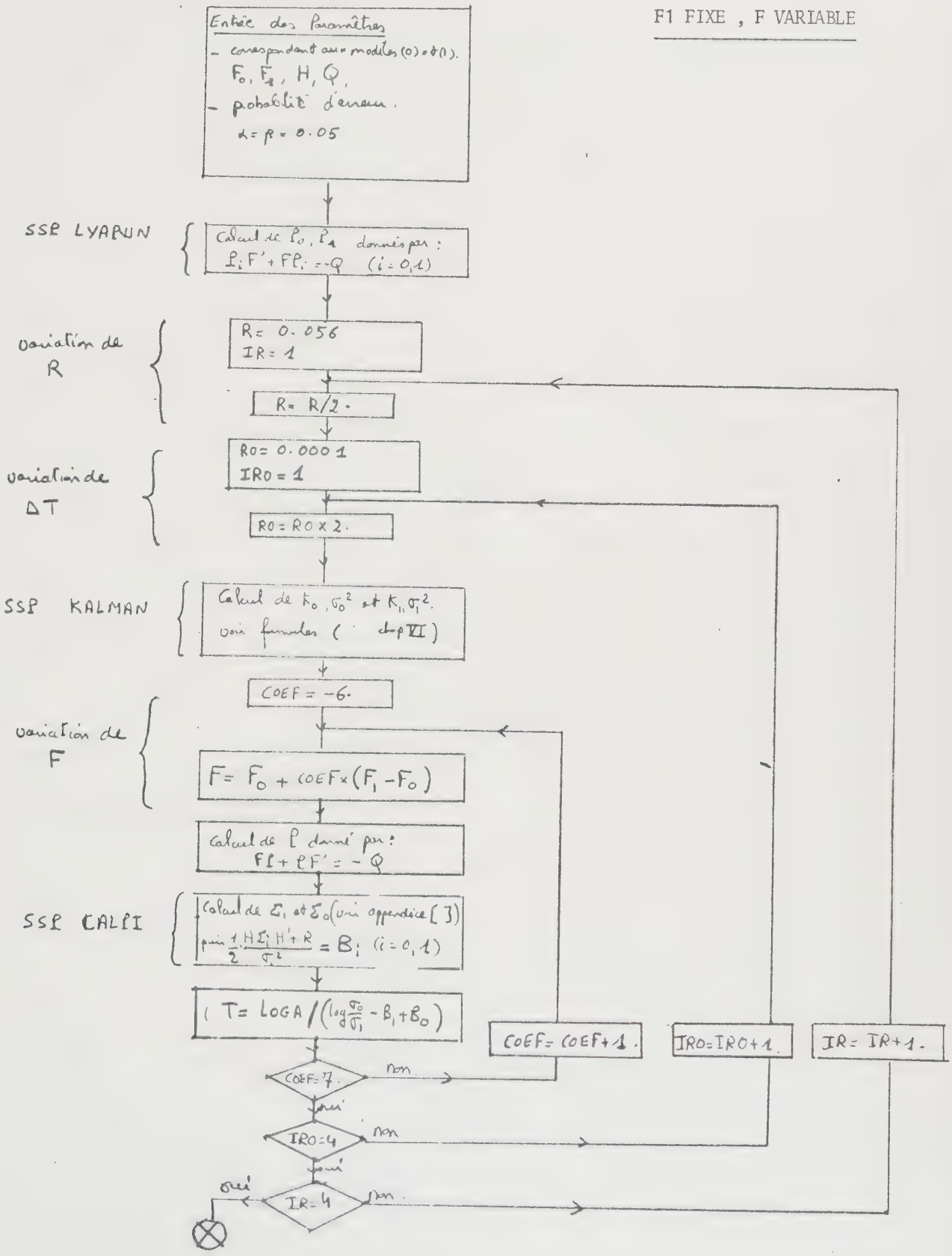

changement
$\alpha_0 = -50$
$\alpha_1 = -40$
$\alpha = -40.$


changement
$\alpha_0 = -50.$
$\alpha_1 = -40.$
$\alpha = -30.$

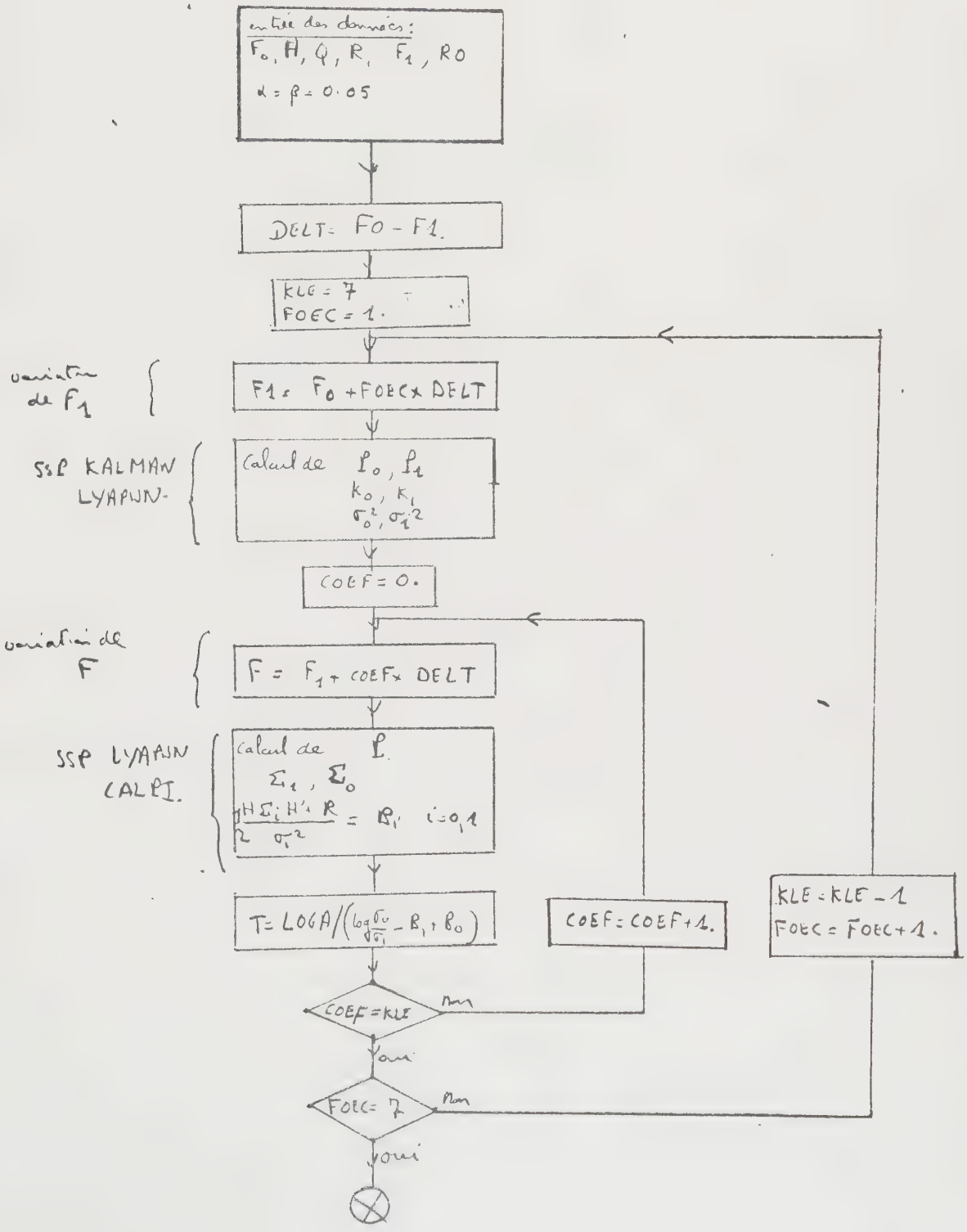
ANNEXE 4

ORGANIGRAMME DE CALCUL DE $T = \log A / E_0 \theta$ AVEC R ET T VARIABLES

F1 FIXE , F VARIABLE



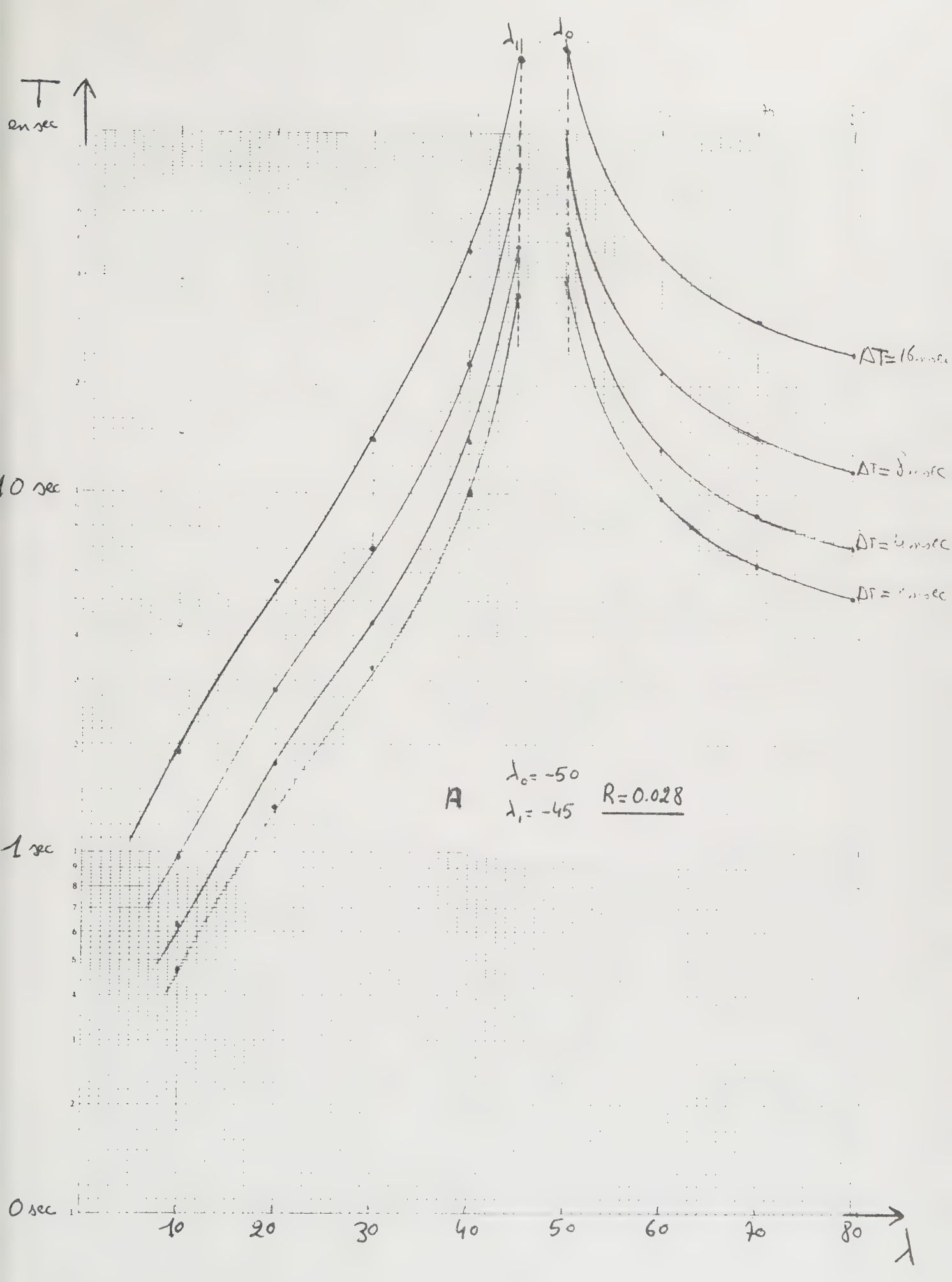
ORGANIGRAMME DE CALCUL DE $T = \log A / E_f \theta$ AVEC F1 VARIABLE ET R ET T FIXES

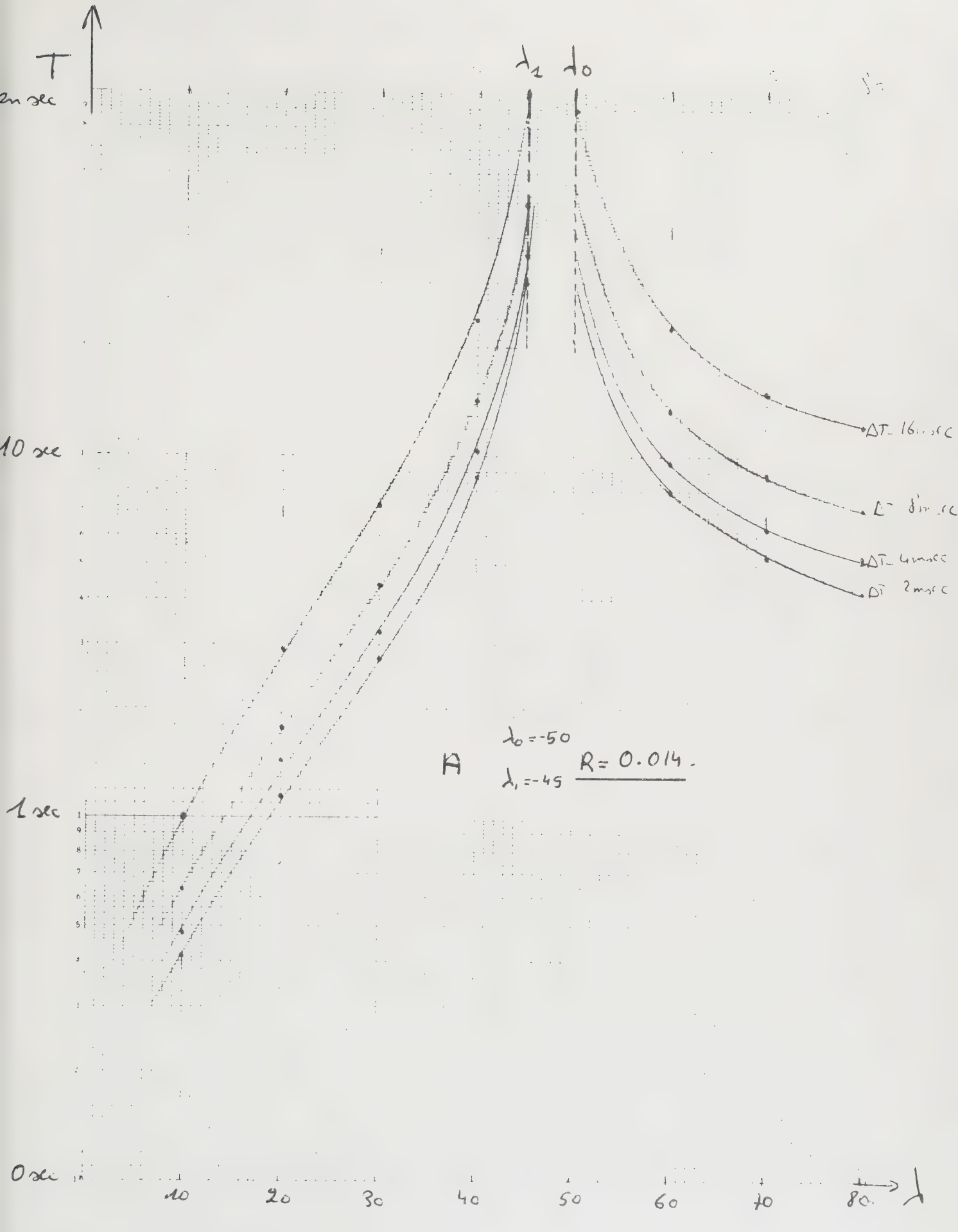


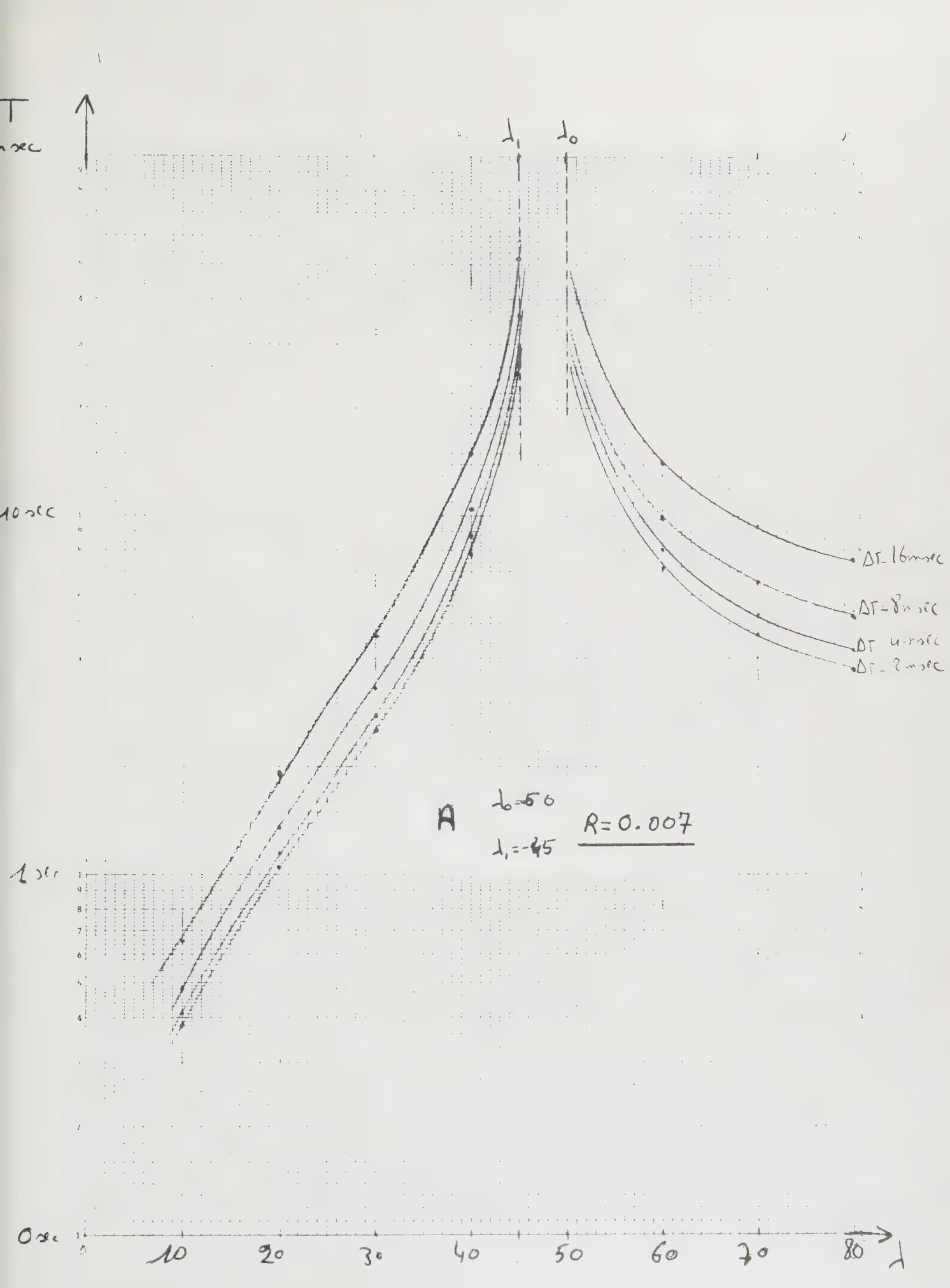
ANNEXE 5

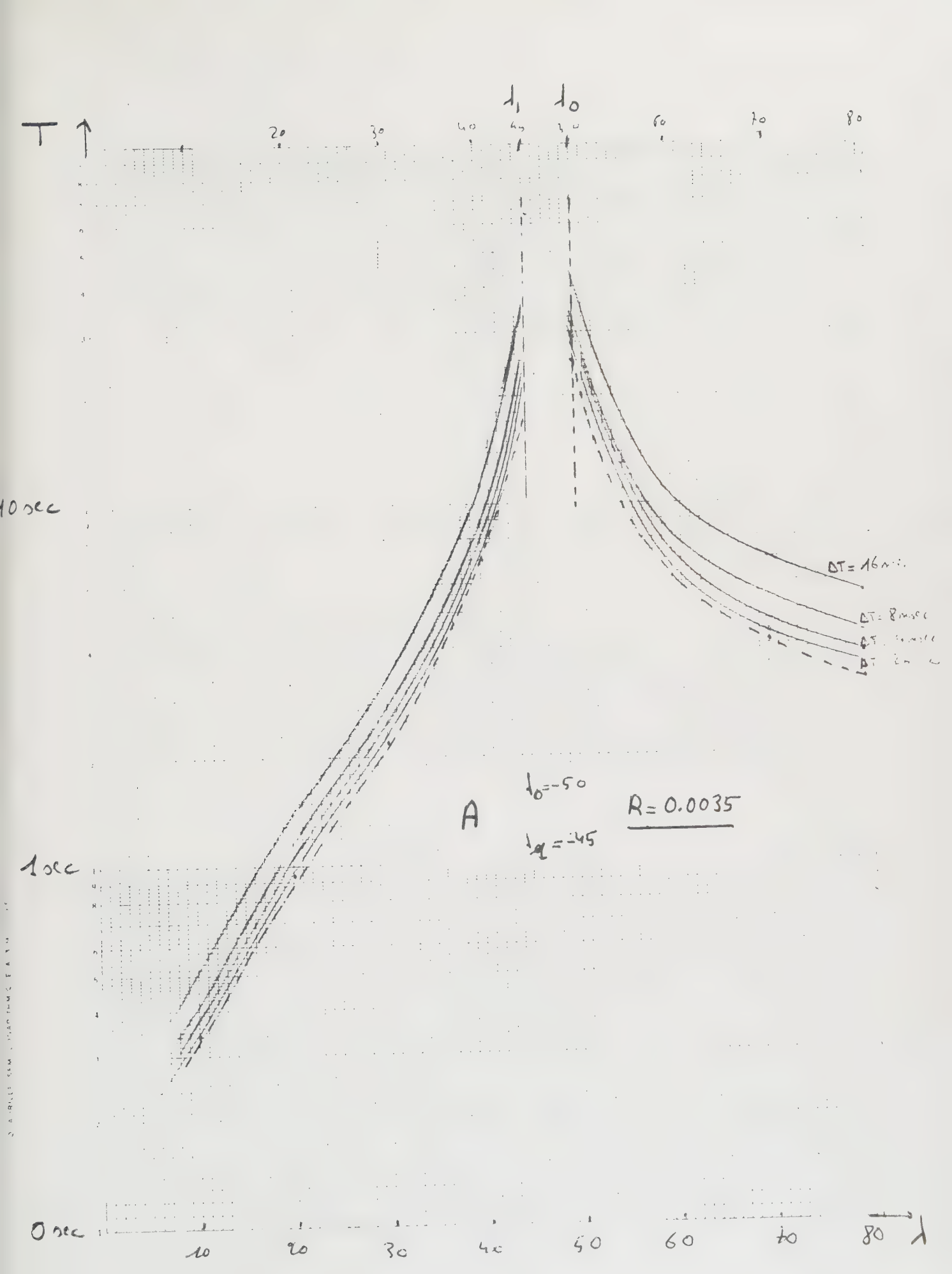
- Temps d'observation T , la vraie valeur de F pouvant être ni Fo, ni F_1
  - système d'ordre 1 , courbes A
  - système d'ordre 2 , courbes B

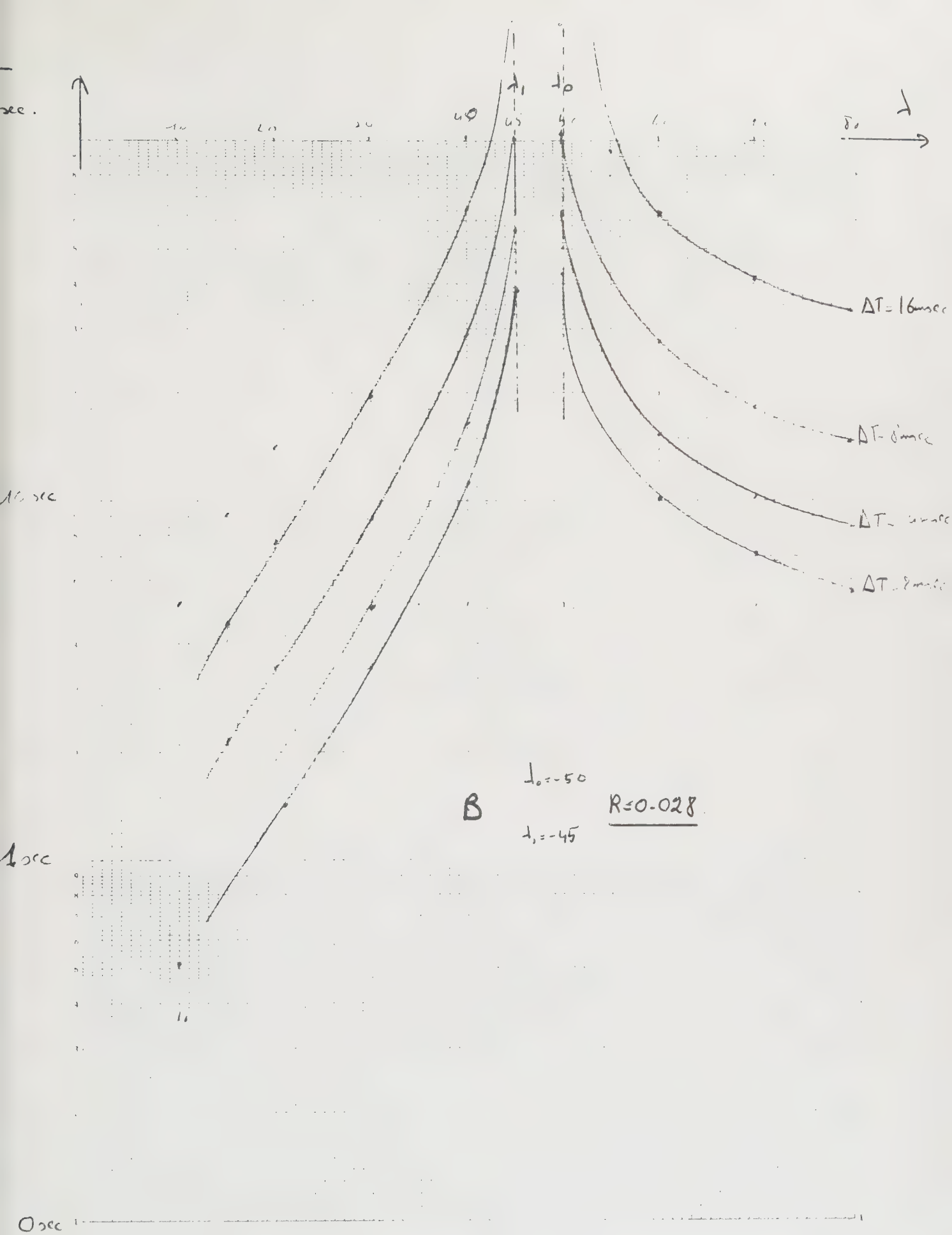
- Sensibilité de T aux variations de F_1
  - système d'ordre 1 , courbes C
  - système d'ordre 2 , courbes D


λ_1
λ_0
T
en sec
10 sec
1 sec
0 sec
ΔT= 16 msec
ΔT= 8 msec
ΔT= 4 msec
ΔT= 2 msec
A
λ_0 = -50
λ_1 = -45
R = 0.028
10
20
30
40
50
60
70
80
λ


T
en sec
λ_1 λ_0
10 sec
1 sec
0 sec
$\lambda_0 = -50$
$\lambda_1 = -45$
R = 0.014.
10
20
30
40
50
60
70
80
λ

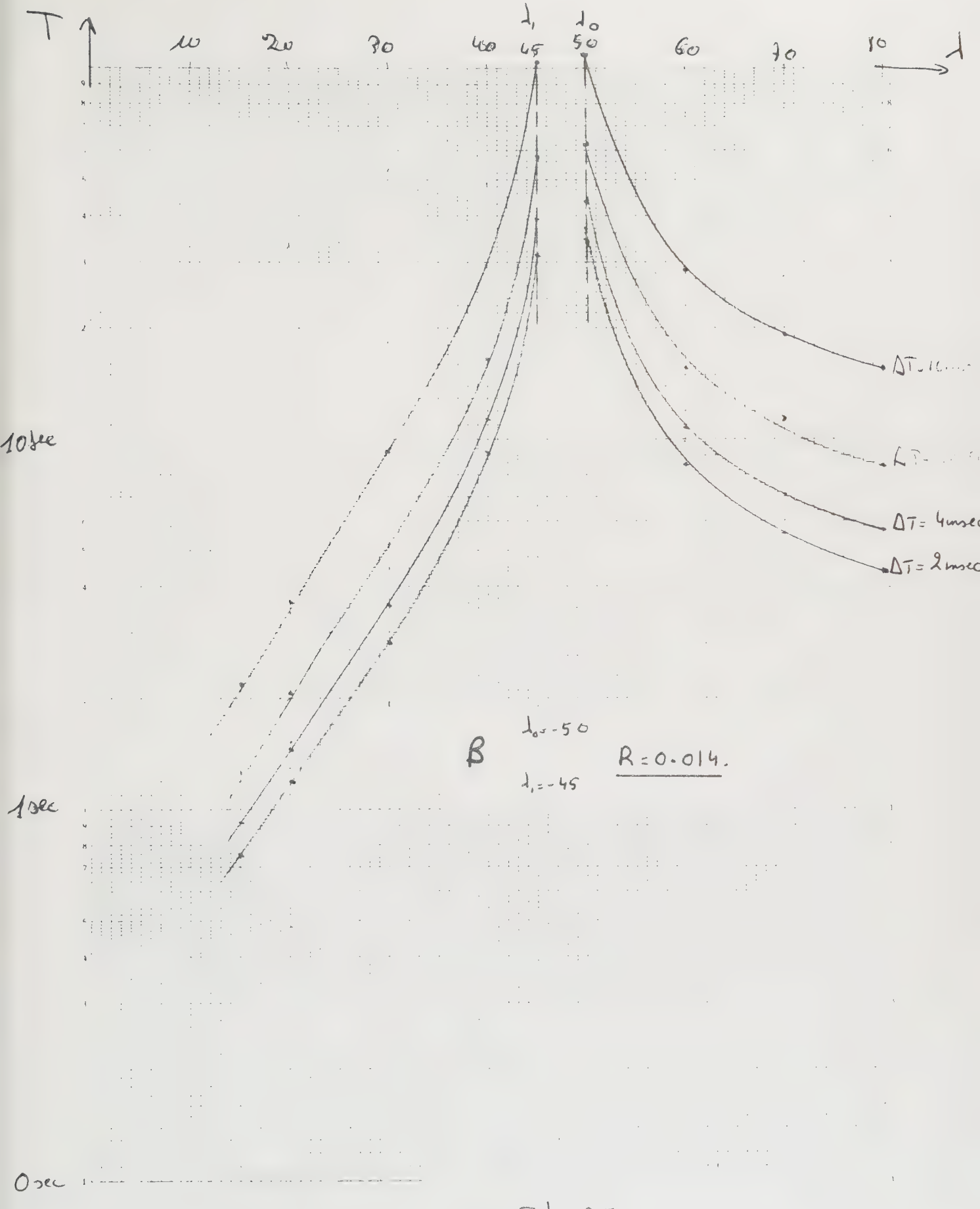

T
msec
λ1
λ0
10 sec
ΔT=16msec
ΔT=8msec
ΔT 4msec
ΔT=2msec
A
λ0=50
λ1=45
R=0.007
1 sec
0sec
10
20
30
40
50
60
70
80
λ


T
λ_1 λ_0
20
30
40
50
60
70
80
10 sec
1 sec
0 sec
ΔT = 16 msec
ΔT = 8 msec
A
λ_0 = -50
λ_1 = -45
R = 0.0035
10
20
30
40
50
60
70
80
λ



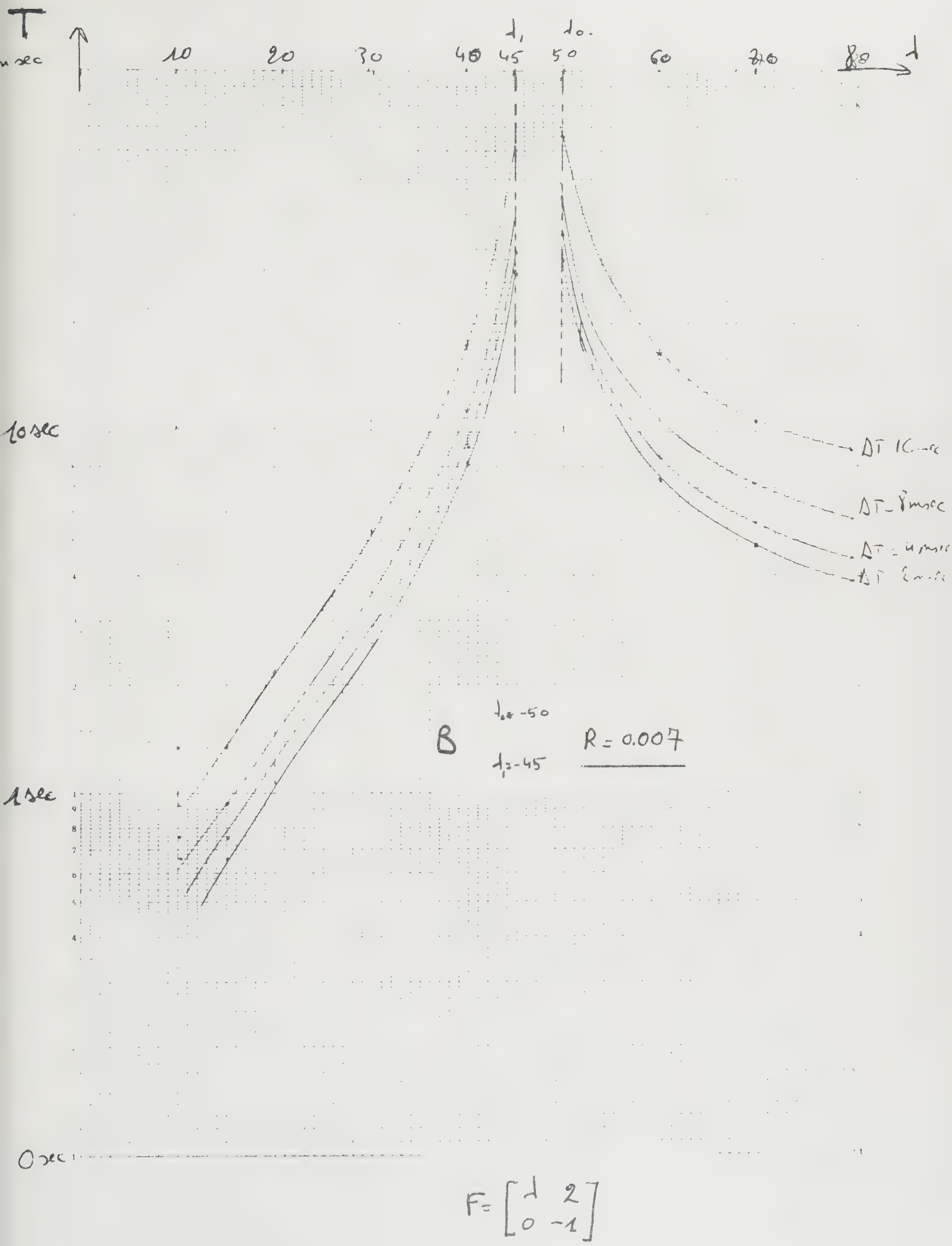
B $\lambda_0 = -50$ $\lambda_1 = -45$ R=0.028

$$F = \begin{bmatrix} \lambda & 2 \\ 0 & -1 \end{bmatrix}$$

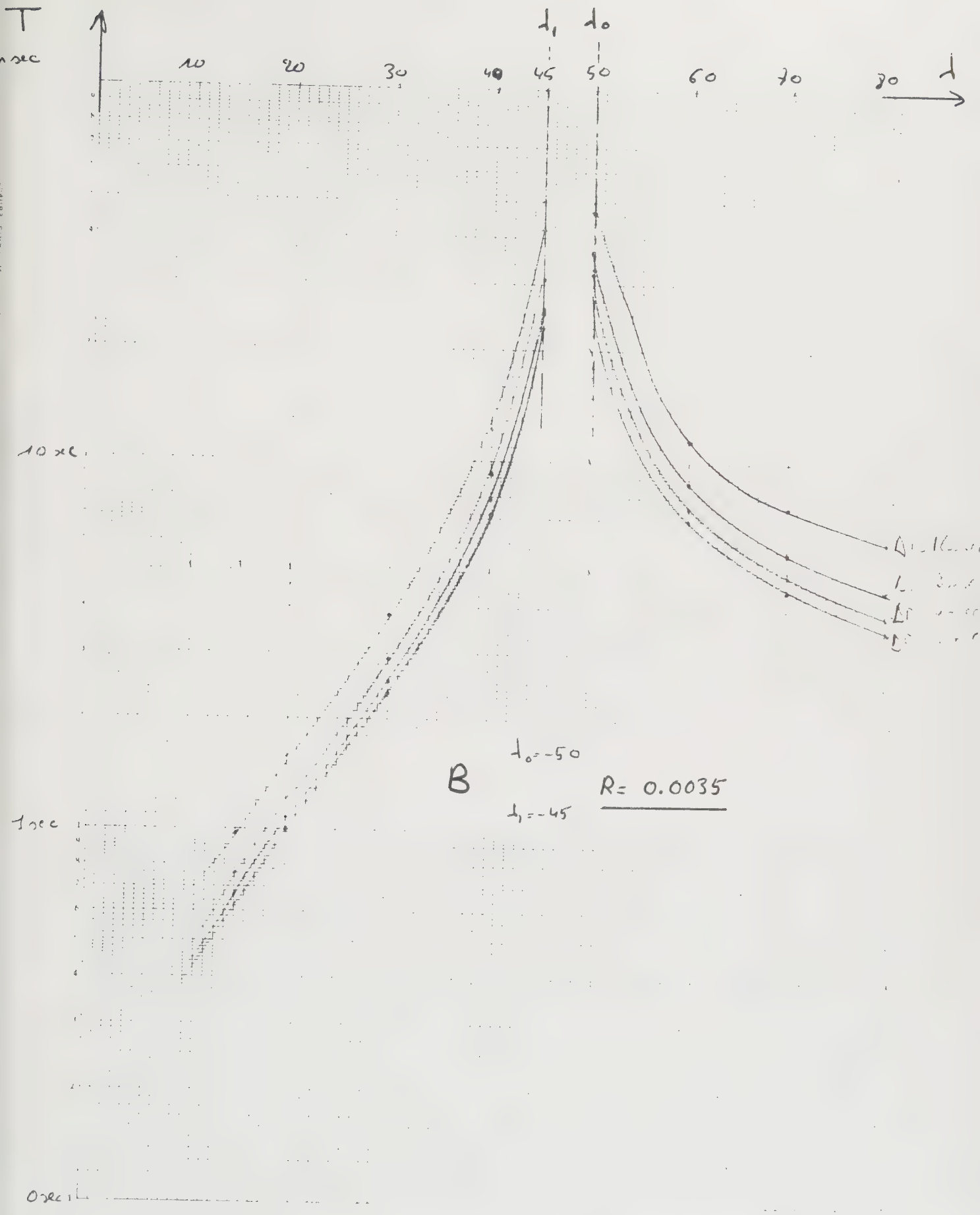


B $\lambda_0 = -50$ $\lambda_1 = -45$ $\underline{R = 0.014.}$

$$F = \begin{bmatrix} \lambda & 2 \\ 0 & -1 \end{bmatrix}$$



$$F = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$$



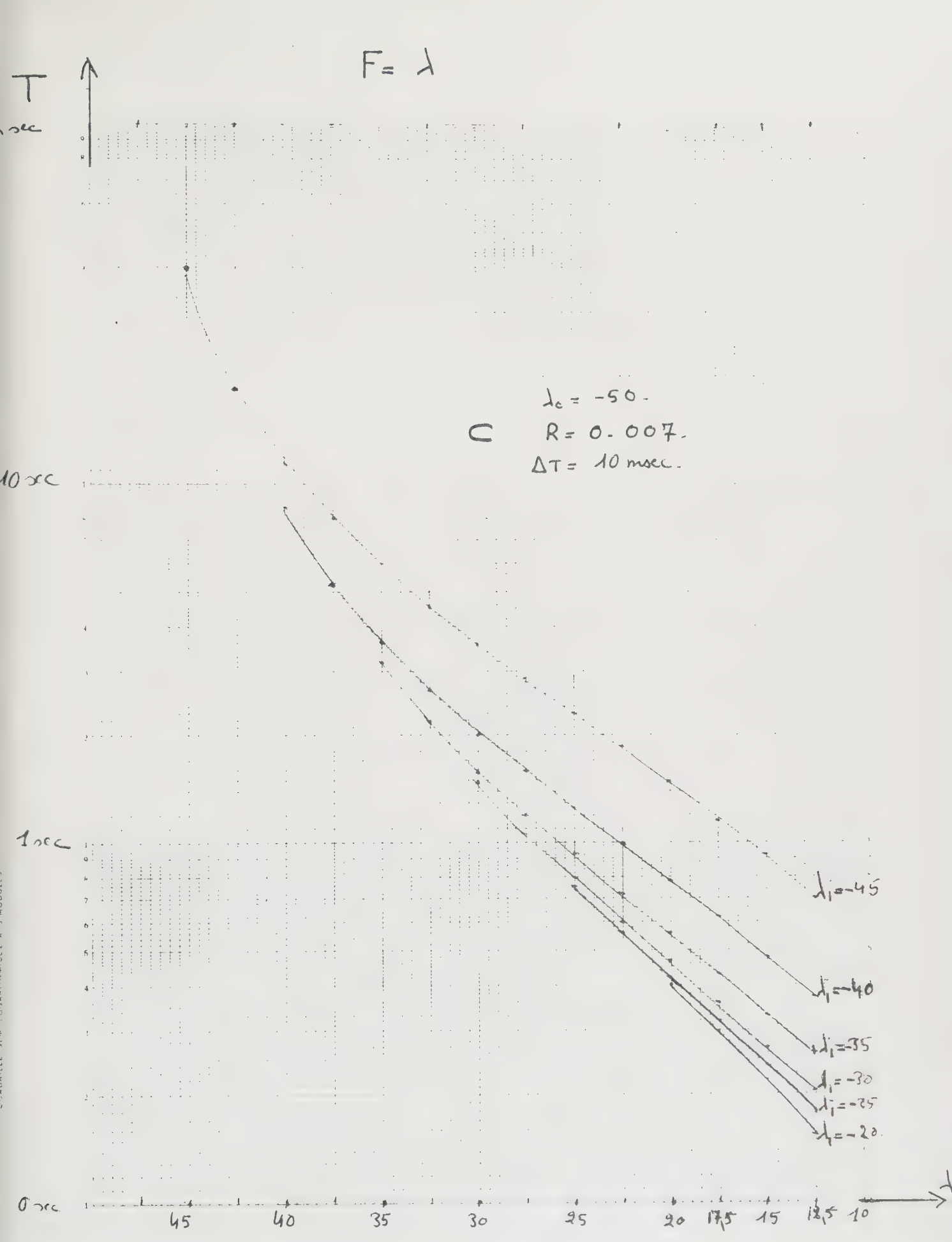
B

$\lambda_0 = -50$

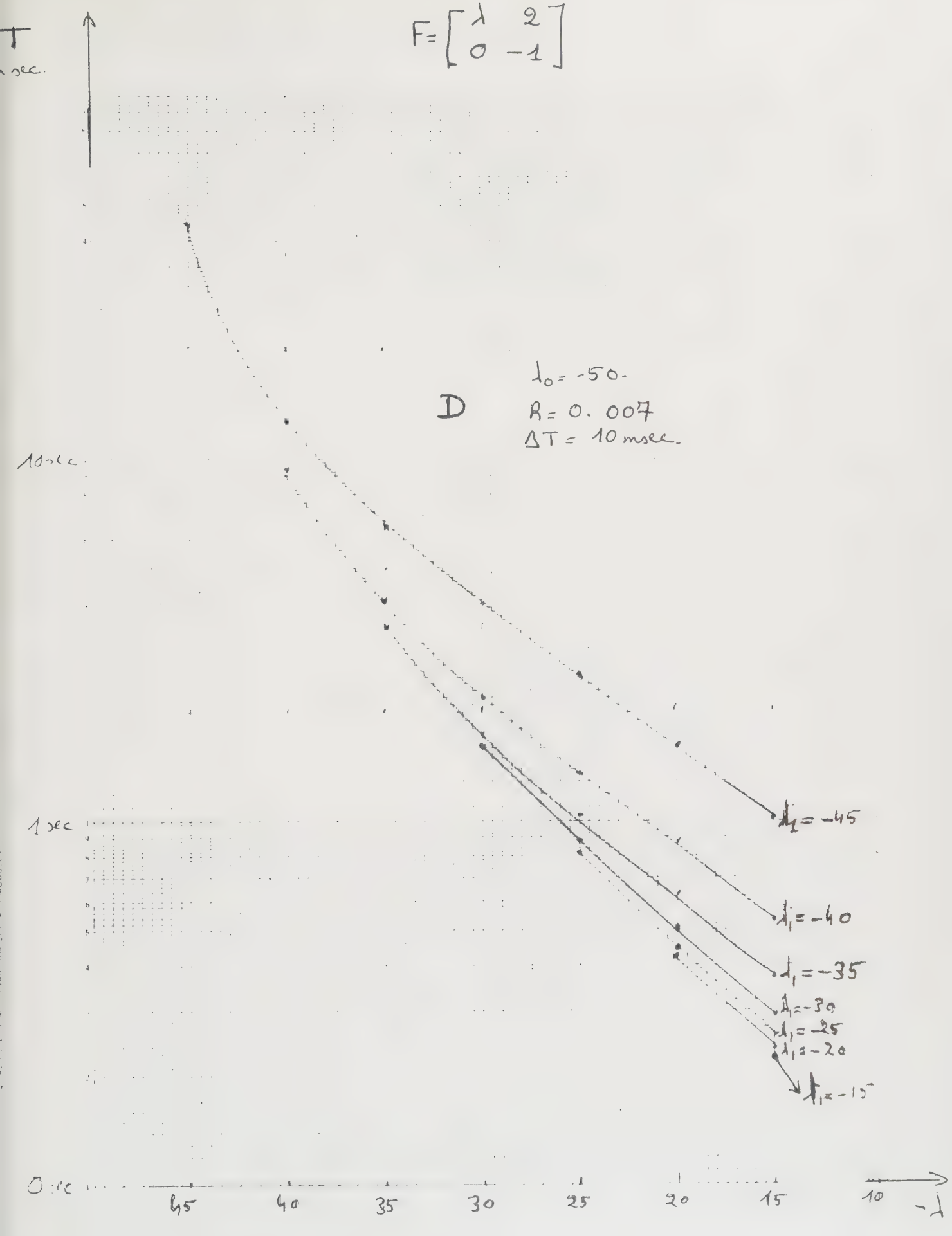
$\lambda_1 = -45$

$R = 0.0035$

$$F = \begin{bmatrix} \lambda & 2 \\ 0 & -1 \end{bmatrix}$$


F= λ
T
sec
10 sec
1 sec
0 sec
C
λc = -50.
R = 0.007.
ΔT = 10 msec.
λ1=-45
λ1=-40
λ1=-35
λ1=-30
λ1=-25
λ1=-20
45
40
35
30
25
20
17,5
15
12,5
10

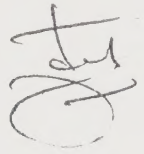
$$F = \begin{bmatrix} \lambda & 2 \\ 0 & -1 \end{bmatrix}$$



DEUXIÈME THÈSE

Propositions données

Vu et approuvé :

Orsay, le 7 décembre 19 73

Le Vice-Président du Centre d'Orsay :